Two Essays on Incomplete Markets

von Diplom-Mathematiker

Felix Esche

aus Berlin

Von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
— Dr. rer. nat. —

genehmigte Dissertation

Vorsitzender des Promotionsausschusses:
Prof. Dr. Günter Frank – Technische Universität Berlin

Berichter:
Prof. Dr. Martin Schweizer – Eidgenössische Technische Hochschule Zürich
Prof. Dr. Alexander Schied – Technische Universität Berlin

Tag der wissenschaftlichen Aussprache: 8. Januar 2004

Berlin 2004
D 83

Abstract

In this thesis we study two aspects of incomplete models in mathematical finance. The first topic is the preservation of the Lévy property of a stochastic process under the minimal entropy martingale measure, for which we give a detailed analysis including a careful explanation of the phenomenon. The second part deals with the approximation of a continuous-time model and the price process of a recursive contingent claim in this model by discrete-time models and price processes, respectively.

Minimization of relative entropy over equivalent martingale measures is the dual problem of maximizing expected exponential utility from terminal wealth over self-financing strategies. We study the relative entropy minimization problem and show that a Lévy process keeps its property of independent and stationary increments under the martingale measure which has minimal relative entropy. The proofs use semimartingale characteristics and Girsanov's theorem for semimartingales. We present a converse of Girsanov's theorem and use this to construct in an explicit way for a given martingale measure with finite relative entropy another martingale measure which reduces the relative entropy and preserves the Lévy property. This uses a representation of relative entropy as a convex functional of two parameters which originate from Girsanov's theorem. This approach enables us to explain why we can identify the entropy-minimizing martingale measure as the Esscher martingale measure. Several applications of these results in mathematical finance are given.

The second part is concerned with the approximation of a continuous-time model by a sequence of finite and discrete-time models. The continuous-time model under consideration is driven by Brownian motions and incorporates non-tradable factors of risk driven by a multivariate point process. Using convergence results for stochastic processes and stochastic differential equations, we first construct a sequence of discrete-time models which converge to the continuous-time model. We then investigate the approximation of the price process for a recursive payoff structure which depends on its own price process. A computation scheme for the corresponding discrete-time price processes gives rise to a representation which corresponds to the pricing approach in continuous time by means of partial differential equations. The convergence of the discrete-time price processes is proved by first showing tightness of the sequence of distributions and then identifying every cluster point of the sequence with the distribution of the continuous-time price process. To establish this, we prove a convergence result for distributions of backward stochastic differential equations.

Zusammenfassung

In dieser Dissertation untersuchen wir zwei Aspekte in der Theorie unvollständiger Märkte. Das erste Thema ist die Erhaltung der Lévy-Eigenschaft eines stochastischen Prozesses unter dem entropieminimierenden Martingalmaß. Wir liefern eine detaillierte Analyse und sorgfältige Erklärung für dieses Phänomen. Der zweite Teil behandelt die Approximation eines zeitstetigen Modells und des Preisprozesses einer rekursiven Auszahlungsstruktur durch zeitdiskrete Modelle bzw. Preisprozesse.

Die Minimierung relativer Entropie über äquivalente Martingalmaße ist das duale Problem zur Maximierung des erwarteten exponentiellen Nutzens aus dem Endvermögen über selbstfinanzierende Handelsstrategien. Wir studieren das Problem der Entropieminimierung und zeigen, dass ein Lévy-Prozess die Eigenschaft unabhängiger und stationärer Zuwächse unter dem Martingalmaß mit minimaler Entropie behält. Die Argumentation benutzt Semimartingalcharakteristiken und den Satz von Girsanov für Semimartingale. Wir stellen eine Umkehrung des Satzes von Girsanov vor, mit der wir in expliziter Weise aus einem gegebenen Martingalmaß mit endlicher relativer Entropie ein neues Martingalmaß konstruieren, das geringere Entropie besitzt und die Lévy-Eigenschaft erhält. Dazu benutzen wir eine Darstellung der relativen Entropie als ein konvexes Funktional zweier Parameter, die durch den Satz von Girsanov gegeben sind. Dieses Vorgehen macht klar, warum wir das entropieminimierende Martingalmaß als das Esscher-Martingalmaß identifizieren können. Zudem werden einige Anwendungen dieser Ergebnisse in der Finanzmathematik untersucht.

Der zweite Teil beschäftigt sich mit der Approximation eines zeitstetigen Modells durch eine Folge von endlichen zeitdiskreten Modellen. Der treibende stochastische Prozess des zeitstetigen Modells im vorliegenden Fall ist eine mehrdimensionale Brownsche Bewegung; zudem beinhaltet es nicht handelbare Risikofaktoren, die von einem multivariaten Punktprozess getrieben werden. Mit Hilfe von Resultaten über die Konvergenz stochastischer Prozesse und stochastischer Differentialgleichungen konstruieren wir eine Folge zeitdiskreter Modelle, die gegen das zeitstetige Modell konvergiert. Danach untersuchen wir die Approximation des Preisprozesses einer rekursiven Auszahlungsstruktur, die neben den handelbaren Anlagen und den nicht handelbaren Risikofaktoren auch von ihrem eigenen Preisprozess abhängt. Zuerst entwickeln wir ein Rechenschema für den zeitdiskreten Preisprozess und erhalten so eine Darstellung, die dem Bewertungsansatz mit partiellen Differentialgleichungen im zeitstetigen Fall entspricht. Wir beweisen die Konvergenz der zeitdiskreten Preisprozesse, indem wir zunächst die Straffheit der Folge der Verteilungen zeigen und dann jeden Häufungspunkt der Folge mit der Verteilung des zeitstetigen Preisprozesses identifizieren. Um das zu erreichen, zeigen wir ein Konvergenzresultat für Verteilungen von Lösungen stochastischer Rückwärtsgleichungen.

Contents

Introduction

Incomplete models have been widely used in mathematical finance because they allow for more realistic modeling than complete models. However they have the drawback that pricing and hedging is not as simple as in the complete case, where every contingent claim is perfectly replicable (and thus redundant), and where the fair price of a contingent claim is given by no-arbitrage arguments alone. In a complete model there exists by the so-called second fundamental theorem of asset pricing a unique martingale measure which is automatically also the pricing measure. In the incomplete case contingent claims are in general not redundant, and one generally faces a wide range of martingale measures so that one must decide which one to use for pricing. As a consequence, one needs to specify hedging preferences such as risk-minimization or utility optimization.

In the celebrated model of Black and Scholes (1973), the stock price is modeled by a geometric Brownian motion, and its completeness is due to the strong property of predictable representation of Brownian motion. Loosely speaking incompleteness of a model occurs if there are more sources of risk than there are tradable assets to hedge away these risks. One possibility to generalize the Black-Scholes model of a one-dimensional Brownian setting is, for instance, to let the volatility of the stock be driven by an additional Brownian motion; this is often referred to as a stochastic volatility model. Another approach is to consider the case where the model is driven by a general semimartingale. The Brownian setting in the former approach seems to be very restrictive, in that it allows jumps neither for the stock price process nor in the volatility. In the latter approach many results on mathematical finance can be formulated and proven, but for explicit results it is too general.

In this thesis we investigate certain aspects of two different models where the incompleteness results from different sources, which may both be viewed as intermediate models to the above formulated "extremes" of incomplete models. In our first model incompleteness comes from the fact that the driving processes are Lévy processes rather than simply Brownian motion, and we show that under the martingale measure with minimal relative entropy among all martingale measures the driving processes remain Lévy processes. This model allows for jumps of the stock price processes as well as stochastic volatility. The second model is situated in a Brownian setting, and the incompleteness stems from untradable factors of risk. Here we are not concerned with the choice of a martingale measure which is optimal in some sense, but we show how such a model and a certain pricing rule in continuous time can be approximated by discrete-time models and pricing rules in the sense of convergence in distribution. This is useful for practical implementations which are usually obtained by a discretization of a continuous-time model.

Minimizing Relative Entropy, and the Lévy Property

The notion of relative entropy of a probability measure with respect to another probability measure, also called Kullback-Leibler information number or I-divergence, and its minimization over a convex set of measures originates from information theory, see, e.g., Kullback (1959). Csiszár (1975) used a more "geometric" approach to handle the problem of minimizing relative entropy over a convex set. It has been used in mathematical finance lately due to its relation with maximizing expected utility in the case of an exponential utility function. Roughly speaking, in a financial market model with price process S maximizing expected exponential utility over a set of admissible trading strategies is dual to finding the entropy-minimizing martingale measure for S. See for instance Delbaen, Grandits, Rheinländer, Samperi, Schweizer and Stricker (2002) or the references given there for this duality and Grandits and Rheinländer (2002) for necessary and sufficient conditions on a candidate measure to have minimal entropy. Both articles treat the case of locally bounded semimartingales. When translated to the Lévy case this calls for bounded jumps, but it turns out that this restriction which is not needed for our results. Note, however, that for a different utility function we get the dual problem of minimizing a different convex functional; see Schachermayer (2001) and Bellini and Frittelli (2002) for a detailed analysis of this issue.

In Part II we are concerned with the following question. Let L be a multi-dimensional Lévy process on an infinite time horizon under some measure P and consider the set of locally absolutely continuous local martingale measures for L. (To include an application in a model with stochastic volatility, we actually consider martingale measures for UL, where U is a linear transformation.) If we pass to a martingale measure Q for L it is well known that L remains a semimartingale under Q, but it is also clear that in general L loses the property of stationary and independent increments. From the mathematical viewpoint it is a natural question to ask whether under some optimality criterion on the new measure the P-Lévy property of L is preserved, and by the duality of minimal entropy with maximal expected exponential utility, this optimality criterion should be minimal entropy if we think of exponential utility. Notice the dependence of relative entropy on the chosen time horizon. If we fix a measure $Q \ll P$ for a *finite* time horizon $[0, T]$ relative entropy of Q with respect to P is an increasing function in T whereas in an *infinite* time horizon setting with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and $Q \overset{\text{loc}}{\ll} P$, the relative entropy of Q with respect to P is in general infinite. We therefore define the *entropy process* at time t of a measure Q which is locally absolutely continuous with respect to P by the relative entropy of the restriction of Q on $\mathcal{F}_t$ with respect to P restricted to $\mathcal{F}_t$, and we say that Q^E is entropy-minimizing if its entropy process is minimal for all $t \geq 0$.

The Lévy property under some measure P of a semimartingale L can be formulated by using the fact that L has independent and stationary increments under P if and only if the P-characteristics of L are deterministic and linear in time. By Girsanov's theorem the change of characteristics of L under a locally absolutely continuous change of measure can be expressed via two parameters β and Y, called the *Girsanov quantities*, which document the change of drift and jump intensities of L when passing to the new measure Q. In general these parameters are stochastic processes and in particular random, so that the Q-characteristics of L are not deterministic any more, and L loses the property of independent stationary increments. On the other hand if β and Y are deterministic and time-independent, then the Q-characteristics of L are deterministic and linear in time, and thus L remains a Lévy process. Moreover a converse of Girsanov's theorem shows that under certain integrability conditions, suitably chosen deterministic and time-independent Girsanov quantities β and Y ensure the existence of a measure with these parameters due to the construction of a nonnegative martingale which serves as density process.

In a filtration generated by a Lévy process, we even have that the density process of Q with respect to P can be expressed by β and Y, which leads us to the result that the relative entropy of Q with respect to P is a convex functional of β and Y. Then by an application of Jensen's inequality and the converse of Girsanov's theorem, relative entropy of Q with respect to P can be reduced if we pass to deterministic and time-independent Girsanov quantities by averaging β and Y over ω and t. Furthermore the local martingale property of L under the new measure Q is characterized by a linear constraint between β and Y, which is preserved under this averaging. From this line of argumentation it is intuitively clear that the martingale measure with minimal entropy, if it exists, must preserve the Lévy property of L.

The insight that the entropy-optimal martingale measure should be sought for in the class of martingale measures for L which preserve the Lévy property of L and the linear constraint between β and Y simplifies the identification of the entropy-optimal martingale measure. In fact, the argumentation leads to a variational problem of minimizing a convex functional over a class of deterministic functions Y, which shows that the so-called Esscher martingale measure for L is the proper candidate for the entropy-minimizing martingale measure.

We provide two different incomplete financial market models where the Esscher martingale measure for the driving Lévy process is indeed the entropy-minimizing martingale measure. The first is a generalization of the Black-Scholes model, where we let $L = (L^i)_{i\in\{1,\dots,d\}}$ be a d-dimensional Lévy process and examine a model $X = (X^i)_{i\in\{1,\dots,d\}}$ of d (discounted) stocks where each X^i is given as the stochastic exponential of L^i. The above results then immediately yield that L is a Lévy process under the entropy-optimal martingale measure for X, and this optimal measure coincides with the Esscher measure for L. The second example is a one-dimensional model X with stochastic volatility in the sense that we are given a two-dimensional (correlated) Lévy process L whose first coordinate drives the stochastic differential equation of the stock price process X, whereas the second coordinate enters in the volatility of X. Here we are interested in the entropy-optimal martingale measure for the first coordinate only, and we show that the optimal measure is given by the Esscher martingale measure for

L^1. However the evolution of L^2 is also modified under the change of measure due to the correlation of L^1 and L^2.

The use of Esscher measures in mathematical finance is not new. Originally Esscher transforms were used in actuarial mathematics, e.g. to calculate a stop-loss premium, but recently they have found their way into mathematical finance as well; see Gerber and Shiu (1994, 1996) for an overview and Bühlmann, Delbaen, Embrechts and Shiryaev (1996) for numerous examples in both discrete and continuous time.

There have been a number of approaches to the problem of minimal entropy in the case of Lévy processes, the most advanced being Fujiwara and Miyahara (2003). They generalize earlier work of Chan (1999), Xia and Yan (2000), Miyahara (2001) and Goll and Rüschendorf (2001) who all impose certain a-priori integrability conditions on the jump measure of L under P. Fujiwara and Miyahara (2003) prove in their Theorem 3.1 the preservation of the Lévy property under the minimal entropy martingale measure for a one-dimensional Lévy process L on a finite time horizon and also identify the optimal measure with an Esscher measure. One drawback in Fujiwara and Miyahara (2003) is that certain integrability issues are not entirely clear from their presentation and that it is not clear how their approach extends to our setup of several dimensions and an infinite time horizon. In addition Fujiwara and Miyahara (2003) merely define the optimal measure via its density and then show that its entropy is minimal, but there is no explanation where the structure of the optimal measure actually comes from and why it preserves the Lévy property.

The idea to express the relative entropy of Q with respect to P via its Girsanov quantities and to minimize the corresponding functional over these parameters already appears in Chan (1999); but there the argument that it is enough to minimize over *deterministic* Girsanov quantities (i.e. over measures under which L remains a Lévy process) is rather heuristic. The argument that the entropy-optimal martingale measure for L must lie in the class of martingale measures which preserve the Lévy property is very similar to a result of Foldes (1990, 1991) who considers an investment problem with market returns given by a process R with independent increments. He proves that an optimal portfolio plan for exponential utility can already be found in the class of deterministic strategies (and is even time-independent if R has independent and stationary increments). Like in our problem, the main techniques used were computations based on semimartingale characteristics.

On the primal side of the problem (i.e. maximizing expected exponential utility), Kallsen (2000) computes for different utility functions the optimal investment strategies in a model of exponential Lévy processes. He also uses semimartingale characteristics, but his line of argumentation is in its nature different from ours.

Convergence of Prices in an Incomplete Model

In order to implement certain models and pricing rules, continuous-time models are often too complex to handle. Therefore it is convenient to both discretize time and space and show that the discretization is good in the sense that the discretized models and pricing rules converge to the continuous-time model as the discretization steps tend to zero. The second subject of this thesis is presented in Part III and treats this aspect of mathematical finance. We study both the approximation of a continuous-time model by a sequence of discrete-time models and the convergence of price processes of a contingent claim. Model convergence then should take place under the real world measures, whereas for the convergence of prices or more generally price processes, one needs to approximate the models under martingale measures. In an incomplete model the price process is not unique, so naturally one should ask *which* price process is under consideration. We do not discuss this question, but rather choose one martingale measure under which we show the desired convergence; this method is known as *pure pricing*. The choice of martingale measure is rather natural, and it can be shown that it coincides in fact with the minimal martingale measure or, under further regularity conditions, the entropy-minimizing martingale measure.

The continuous-time model is taken from Becherer and Schweizer (2003) and consists of a continuous multidimensional process S and a process η with values in a finite set. The process S is the solution of a stochastic differential equation which is driven by a Brownian motion W and where in addition the coefficients depend on η, which is driven by a point process N whose intensities in turn depend on S. The process η is not tradable and therefore accounts for the incompleteness of the market. If we consider S to consist of assets, then we can think of η as the rating given by some agency, which influences the value of the asset and which is itself influenced by its performance.

The existence of such a model with mutual dependences is shown with the help of a change of measure where one first considers the case where under some measure, N is a multivariate standard Poisson process independent of W, so that η is an autonomous process which can be plugged into the stochastic differential equation for S. Then for a suitably chosen change of measure one obtains a model with the desired properties.

Becherer and Schweizer (2003) consider payoff structures which consist of a payment at expiration time, a continuous flow of payments up to expiration time and an additional lump sum payment every time the process η jumps from one state to another. This payoff structure is not a classical contingent claim since it depends not only on the stock price process and the untradable factors of risk, but also on its own price process. This feature is seen, e.g., in models of defaultable bonds with fractional recovery. By the pure pricing approach Becherer and Schweizer (2003) *define* the price at time t of such a payoff structure as the conditional expectation of all future payments under an a-priori chosen martingale measure for S, given the information up to time t. But since the payments also depend on the price process itself, this “definition” of the price process is at first not well-defined. However if one formalizes the price process V as the solution of a corresponding backward stochastic differential equation, one has the necessary tools to show existence and uniqueness for a solution of such an equation. Moreover, Becherer and Schweizer (2003) show that due to the Markov structure of the driving processes, V is given by $V_t = v(t, S_t, \eta_t)$, where v solves a certain reaction-diffusion equation.

The approximation of the *model* by a sequence of discrete-time models goes along the same lines as the proof of its existence. We first construct a model in discrete time by choosing for each n a stochastic basis with independent binomial processes W^n and N^n which converge in distribution to W and N as n tends to infinity. The model (S^n, η^n) itself is then given as the solution of stochastic difference equations corresponding to the stochastic differential equations from the continuous-time model. The convergence of (S^n, η^n) to (S, η) is shown by applying results by Kurtz and Protter (1991), Duffie and Protter (1992) and Kurtz and Protter (1996) on the convergence of solutions of stochastic differential equations for converging driving processes. For suitably chosen changes of measure this convergence is then preserved under a sequence of contiguous measures which allow for the desired mutual dependences in the discrete-time models.

Concerning the approximation of *price processes* the reasoning is less straightforward. For a discretized analogue of the above payoff structures we need to show convergence of solutions of a sequence of backward stochastic difference equations. In order to achieve this we give a backward computation scheme for the price process V^n and we show in analogy to the continuous-time case that $V^n_t = v^n(t, S^n_t, \eta^n_t)$, where v^n is sufficiently smooth. Instead of showing convergence of the sequence $(v^n)_{n\in\mathbb{N}}$ of functions to the solution v of a reaction-diffusion equation (which looks delicate), we opt for the classical approach to show convergence of a sequence of semimartingales, namely to first show tightness of the sequence of distributions and then to identify every cluster point of the sequence with the same limit. Due to the representation of V^n via the functions v^n and the regularity properties of v^n, the first point is, though lengthy, relatively straightforward, using a general tightness criterion by Jacod, Mémin and Métivier (1983). However, the second point is more delicate since the processes V^n and V via their backward representations crucially depend on the chosen filtrations, whereas convergence in distribution is a property of càdlàg processes and has at first nothing to do with filtrations. This problem is circumvented by using the Skorokhod embedding theorem, so that the tightness of $(V^n)_{n\in\mathbb{N}}$ implies P-a.s. convergence of subsequences of $(V^n)_{n\in\mathbb{N}}$. Moreover, this method allows us to use the concept of *convergence of filtrations*, as developed by Hoover (1991) and extended by Antonelli and Kohatsu-Higa (2000) and Coquet, Mémin and Słominski (2001), in order to identify every cluster point of a subsequence of $(V^n)_{n\in\mathbb{N}}$ with V from the continuous-time model. As a by-product we obtain a method to convert backward equations into forward equations once a candidate solution of the backward equation is given.

The question of approximating models and price processes in mathematical finance dates back to the celebrated paper of Cox, Ross and Rubinstein (1979) who show that in a suitably rescaled binomial model the price of a European option converges to the corresponding price in the Black-Scholes model. In their paper only the behaviour of the given sequence of prices is analyzed without showing convergence of the binomial model nor of the price *processes*. By Donsker's theorem the convergence of a rescaled binomial model to the Black-Scholes model of geometric Brownian motion is a well-known result, whereas convergence of price processes requires more advanced results concerning preservation of model convergence under a change of measure and the above mentioned convergence of filtrations. If one is interested solely in the convergence of *prices* (i.e. convergence of the initial value of the price process at time 0), and if the contingent claim satisfies certain regularity conditions, it is sufficient to show convergence of the models under a sequence of martingale measures. For this issue see Hubalek and Schachermayer (1998) who use results on contiguity of sequences of probability measures which ensure that the convergence of models is preserved when passing to martingale measures. Their models are not necessarily complete, and they give explicit conditions for convergence of prices in discrete-time models of row-wise independent triangular arrays. See also Hubalek and Hudetz (1998) for a result concerning convergence of models (and thus prices) along the sequence of entropy-minimizing martingale measures if the limit model is complete. In the case of an incomplete limit model, where the stock price process is given by a special semimartingale, Prigent (1999) and Lesne, Prigent and Scaillet (2000) show that under further regularity conditions model convergence under the real-world measure implies model convergence along the sequence of minimal martingale measures.

It is a different issue to ask for convergence of *strategies*. In a model where the driving processes possess the strong property of predictable representation, i.e. when the model is complete, the claim can be written as an initial value plus a stochastic integral of the strategy with respect to the driving processes. The unique price of the claim at time t is then given by the initial value plus the value of the integral at time t. In an incomplete model this is not the case any more. But for a given hedging strategy one defines its value process analogously as the initial capital plus the stochastic integral of the strategy with respect to the price process of the asset under consideration. For the case of risk-minimizing strategies and under the assumption that asset prices are martingales under the real world measure, Jacod, Méléard and Protter (2000) show conditions under which joint convergence of models and contingent claims implies convergence of strategies.

The above concepts can only be applied when the payoff structure does not depend on its own price process. If one considers the case where one has such a dependence, one needs to approximate the solution of a backward stochastic differential equation in order to obtain functional convergence of the price processes. There are a number of articles on this subject, starting with Douglas, Ma and Protter (1996), who approximate forward-backward stochastic differential equations by means of their so-called "four step scheme" and the relation of a backward stochastic differential equation with a quasi-linear partial differential equation, as introduced by Ma, Protter and Yong (1994). Ma, Protter, San Martín and Torres (2002) also approximate the solution of a quasi-linear partial differential equation, but it seems that their proof of tightness of the sequence of solutions of the corresponding backward stochastic difference equations lacks clarity and involves an improper use of the tightness criterion of Jacod, Mémin and Métivier (1983) also used in this thesis to show tightness.

There are some results on the approximation of backward stochastic differential equations in the form of Duffie and Epstein (1992), i.e. when the solution process can be written as the solution of an equation which involves conditional expectations as in our context; see Briand, Delyon and Mémin (2001) for a space-time discretization and Bouchard and Touzi (2002) who only discretize time but not space. Both approaches involve the use of the concept of convergence of filtrations but have the natural drawback that all processes need to be defined on the same probability space. In addition, the limiting filtration is generated by Brownian motion only, so that these results cannot be easily extended to our setting of Brownian and Poisson filtration.

Organization of the Thesis

In Part I we collect the most important known results which are needed in the development of this thesis. Section 1.2 deals with the theory of Lévy processes, and in Section 1.3 we recall results on relative entropy. We also define the *entropy process* and certain sets of martingale measures over which we minimize relative entropy in Part II. Sections 1.4 and 1.5 a collection of results on characteristics and convergence of semimartingales and Lévy processes. Most of these results are known and only cited from the existing literature or at most adjusted to our situation. One exception is Theorem 1.34, by which it is possible to express the characteristics of linear transforms of semimartingales, a result which has so far only been stated for the case of stationary and independent increments. Also note Theorem 1.53, which may be viewed as a point process version of a result by Kurtz and Protter (1996) on the convergence of solutions of stochastic differential equations.

Part II is devoted to the development of the above mentioned preservation of the Lévy property under the minimal entropy martingale measure. This is an extended and adapted version of Esche and Schweizer (2003). In Chapter 2 we recall the most important results on changes of measures and density processes in the case where the processes under consideration are Lévy processes under the original measure. In Section 2.1 we recall Girsanov's theorem and introduce the *Girsanov quantities* β and Y which describe the change of drift and jump intensities as mentioned above. We give an example concerning the change of measure in the case of a geometric Lévy process, and in another example we show that the distribution of a Brownian motion with random drift is still determined by its characteristics, a fact which is generally true only for Lévy processes and not for arbitrary semimartingales. Section 2.2 deals with the explicit computation of the density process for a given measure, whereas in Section 2.3 we present a converse of Girsanov's theorem, which is indispensable for the construction of martingale measures with smaller relative entropy than a given martingale measure.

Chapter 3 is the heart of Part II. There we show that a P-Lévy process L remains a Lévy process under the entropy-minimizing martingale measure for UL, where U denotes a fixed linear transformation. In Section 3.1, which is rather technical, we provide an explicit calculation for the relative entropy, and we show how to average the Girsanov quantities of a given measure in order to construct a new measure with smaller entropy. The main result of Section 3.2 is the parametrization of the class of martingale measures for UL in terms of the Girsanov quantities. In Section 3.3 we finally show that the Lévy property is preserved under the entropy-optimal martingale measure. We remark here that the intuitive idea of simply averaging the Girsanov parameters β and Y only works rigorously under additional integrability conditions on the jumps of L. To obtain our result in full generality therefore involves some additional approximation arguments.

In Chapter 4 we provide several applications of the preservation of the Lévy property from Chapter 3. By a variational argument we identify in Section 4.1 the candidate for the entropy-optimal martingale measure as the Esscher martingale measure, and we present in Section 4.2 several properties of Esscher measures including a calculation of their Girsanov quantities. Section 4.3 contains a generalization of the Black-Scholes model where the driving process is a multidimensional Lévy process, and in Section 4.4 we examine the above mentioned one-dimensional Lévy model with stochastic volatility.

Part III addresses the topic of approximating a class of continuous-time models by models with discrete time. In Chapter 5 we present the continuous-time model (S, η) of Becherer and Schweizer (2003) with mutual dependences between stock prices S and untradable factors of risk η and show convergence of a sequence of corresponding discrete-time models. Section 5.1 is concerned with the existence of the continuous-time model with mutual dependences, where we follow the approach of Becherer and Schweizer (2003) in that we first show existence of a model where the driving processes are independent and then obtain the mutual dependences by a suitable change of measure. In Section 5.2 we introduce a sequence of discrete-time binomial models which are constructed to converge in distribution to the continuous-time model in the case of independent driving processes. We obtain in Section 5.3 convergence in the more general situation of mutual dependences by an additional approximation of the density process which is responsible for the change of measure in the continuous-time case. Finally in Section 5.4 we present a detailed analysis of the factors of risk and their influence on the above change of measure.

In Chapter 6 we extend the convergence results to an approximation of the price process V of a generalized contingent claim when V is given as the solution of a backward stochastic differential equation. In Section 6.1 we recall the structure of the payoff as presented in Becherer and Schweizer (2003) and construct a discrete-time analogue V^n of the price process via a backward stochastic difference equation. In Section 6.2 we give a backward computation scheme for V^n which uses the Markov structure of the model and is related to the pricing approach with partial differential equations in the continuous-time case. Section 6.3 is concerned with the tightness of the sequence $(V^n)_{n\in\mathbb{N}}$ of discrete-time price processes, and in Section 6.4 we show that we indeed have convergence of the solutions V^n for the backward stochastic difference equations in the discrete-time models to the solution V for the backward stochastic differential equation.

We have relegated results to the Appendix if they are not directly needed for the development of our results or if their proof is long and technical and would therefore disrupt the flow of reading. In Appendix A we give a short summary of results on infinitely divisible distributions and their relation with Lévy processes. These results allow us to construct equivalent martingale measures which preserve the Lévy property of a Lévy process directly via the characteristic triplet of the corresponding infinitely divisible distribution. In Appendix B we recall the definition of the Skorokhod topology on the Skorokhod space $\mathbb{D}(\mathbb{R}^d)$ of $\mathbb{R}^d$-valued càdlàg functions on $\mathbb{R}_+$. These results are heavily used for the convergence results in Part III, and we prove the continuity of a number of mappings on the Skorokhod space. Several technical results which are needed in this thesis can be found in Appendix C. Appendix D contains a detailed discussion of the article by Fujiwara and Miyahara (2003) on the preservation of the Lévy property under the entropy-optimal martingale measure, and we present a comparison to our results from Part II.

Acknowledgments

Foremost, it is a pleasure to express my deep gratitude to my advisor Martin Schweizer for guiding me through the fascinating field of mathematical finance. Throughout the time I was working on this thesis he always shared with me his wide knowledge and experience of probability theory and mathematical finance, and without his valuable suggestions and comments the realization of this project would not have been possible. I also wish to thank Alexander Schied for readily accepting the task of being co-examiner of my thesis and for the discussions and suggestions on the presentation of several issues before the disputation.

Furthermore I should like to thank my friends and colleagues Stefan Ankirchner, Peter Arts, Peter Bank, Steffen Dereich, Georgi Dimitroff, Matthias an der Heiden, Ulrich Horst, Anis Matoussi, Silke Meiner, Beat Niederhauser, Christian Vormoor, Stefan Weber and Wiebke Wittmüß for fruitful discussions and for generating the pleasant working environment at Technische Universität Berlin I have enjoyed so much over the years. Special thanks go to Dirk Becherer who not only encouraged me to finish my thesis but also never stopped to come up with inspiring questions and invaluable ideas on how to further develop the subjects of this thesis.

Last but not least I deeply thank Anne and Jonathan for all their love, patience and support.

Part I

PRELIMINARIES

Chapter 1

Preliminaries

In this chapter we recall the most important known results needed in this dissertation. In Section 1.1 we give a short summary of frequently used notions and notation. In Sections 1.2–1.4 we present mainly citations from the known literature on Lévy processes, relative entropy and characteristics of semimartingales, which are sometimes adapted to fit into a consistent notation scheme. These results will be needed in Part II. In Section 1.4 we have added some easy applications to characteristics of linear transformations of semimartingales which are later needed when we consider a model with stochastic volatility.

In Section 1.5 we summarize general results on weak convergence of probability measures, tightness of càdlàg processes, convergence in distribution of semimartingales, convergence of solutions of stochastic differential equations, and convergence under changes of measures. We also give several examples which will serve as references in Part III of this dissertation.

1.1 Notions and Notation

We choose any filtration on a measurable space $(\Omega, \mathcal{F})$ to be a *right-continuous* increasing family $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ of sub-σ-fields of $\mathcal{F}$ and by convention we set $\mathcal{F} = \mathcal{F}_\infty = \bigvee_{t\geq 0} \mathcal{F}_t$. For a stochastic process X we denote by $\mathbb{F}^X$ the smallest right-continuous filtration to which X is adapted; more precisely we let $\mathcal{F}_t^{X,0} = \sigma(X_s, s \leq t)$ and define $\mathcal{F}_t^X = \bigcap_{s>t} \mathcal{F}_s^{X,0}$. We then call $\mathbb{F}^X$ the *filtration generated by* X. If we are given a probability measure P on $(\Omega, \mathcal{F})$ and if $\mathbb{F}$ is a filtration, we denote by $\mathbb{F}(P)$ the P-augmentation of $\mathbb{F}$.

Sometimes it will be necessary to work on the path space of càdlàg (i.e. right-continuous with left limits) semimartingales. For $I \subseteq [0, \infty)$ we denote by $\mathbb{D} = \mathbb{D}(I, \mathbb{R}^d)$ the Skorokhod space of all càdlàg functions $\alpha\colon I \to \mathbb{R}^d$. For $I = \mathbb{R}_+ = [0, \infty)$ we denote by $\mathcal{D}_t^0(\mathbb{R}^d)$ the σ-field generated by all mappings $\alpha \mapsto \alpha(s)$ for $s \leq t$, and $\mathcal{D}(\mathbb{R}^d) = \bigvee_{t\geq 0} \mathcal{D}_t^0(\mathbb{R}^d)$, and $\mathcal{D}_t(\mathbb{R}^d) = \cap_{s>t} \mathcal{D}_s^0(\mathbb{R}^d)$. Then $\mathbf{D} = (\mathcal{D}_t(\mathbb{R}^d))_{t\geq 0}$ is a filtration, which we call the canonical filtration on $\mathbb{D}([0, \infty), \mathbb{R}^d)$. For $I = [0, T]$ an analogous concept of filtration applies. If we endow $\mathbb{D}$ with the Skorokhod topology, it is a Polish space, and we denote by $\mathcal{B}(\mathbb{D})$ the Borel σ-algebra on $\mathbb{D}$, which coincides with $\mathcal{D}(\mathbb{R}^d)$ as defined above (cf. Jacod and Shiryaev (1987), Theorem VI.1.14).

On $\Omega \times \mathbb{R}_+$ we denote by $\mathcal{P}$ the predictable σ-field, which is generated by all càg (i.e. left-continuous) adapted processes, and by $\mathcal{O}$ the optional σ-field, which is generated by all càdlàg adapted processes. For the introduction and use of random measures and characteristics of $\mathbb{R}^d$-valued semimartingales it is necessary to define on $\widetilde{\Omega} = \Omega \times \mathbb{R}_+ \times \mathbb{R}^d$ the σ-fields $\widetilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}^d$ and $\widetilde{\mathcal{O}} = \mathcal{O} \otimes \mathcal{B}^d$, where $\mathcal{B}^d$ is the Borel σ-field on $\mathbb{R}^d$. We call a function $W : \widetilde{\Omega} \to \mathbb{R}$ *optional* (respectively *predictable*) if it is $\widetilde{\mathcal{O}}$- (respectively $\widetilde{\mathcal{P}}$-) measurable. Note that $\mathcal{P} \subseteq \mathcal{O}$ and thus $\widetilde{\mathcal{P}} \subseteq \widetilde{\mathcal{O}}$.

Observe the difference between a predictable *process* which is a $\mathcal{P}$-measurable mapping on $\Omega \times \mathbb{R}_+$ and a predictable *function* which is a $\widetilde{\mathcal{P}}$-measurable mapping on $\widetilde{\Omega}$. Also note that every predictable process may be considered as a predictable function which is constant in the last argument.

Let P, Q be two measures on the filtered measurable space $(\Omega, \mathcal{F}, \mathbb{F})$. We say that Q is *locally absolutely continuous* (respectively *locally equivalent*) with respect to P, and we write $Q \overset{\text{loc}}{\ll} P$ (respectively $Q \overset{\text{loc}}{\sim} P$), if $Q|_{\mathcal{F}_t} \ll P|_{\mathcal{F}_t}$ (respectively $Q|_{\mathcal{F}_t} \sim P|_{\mathcal{F}_t}$) for all $t \geq 0$.

We call *stochastic basis* a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$, and we take all semimartingales to have càdlàg paths. If X is a semimartingale on $(\Omega, \mathcal{F}, \mathbb{F}, P)$, we denote by X^c the continuous (P-)martingale part of X (cf. Jacod and Shiryaev (1987), Proposition I.4.27). X^c is a continuous local P-martingale and unique up to indistinguishability.

We call $h : \mathbb{R}^d \to \mathbb{R}^d$ a *truncation function* if h is bounded with compact support and satisfies $h(x) = x$ in a neighbourhood of 0. The truncation function $h_0(x) = x \mathbb{1}_{\{|x| \leq 1\}}$ is called the *canonical truncation function*.

If τ and σ are two stopping times, we define the *stochastic interval* $[\![\tau, \sigma]\!]$ by

$$[\![\tau, \sigma]\!] = \{(\omega, t) \in \Omega \times \mathbb{R}_+ : \tau(\omega) \leq t \leq \sigma(\omega)\};$$

the intervals $[\![\tau, \sigma[\![$, $]\!]\tau, \sigma]\!]$ and $]\!]\tau, \sigma[\![$ are defined analogously.

For a d-dimensional local martingale X for which each component is locally square-integrable we write $\langle\!\langle X \rangle\!\rangle := \sum_{i=1}^d \langle X^i \rangle$.

If A is a process of finite variation we denote its total variation process by $\mathrm{Var}(A) = \int |dA|$.

1.2 Lévy Processes

The concept of processes with stationary independent increments, also called *Lévy processes*, is widely used in mathematical finance to model the driving processes of stochastic differential equations which describe the evolutions of financial assets. On the one hand, such processes are generalizations of Brownian motion in that they allow the processes under consideration to have jumps, whereas on the other hand they are easier to handle as general semimartingales. Here we recall the most important facts on Lévy processes which are needed later on in Part II.

Definition 1.1 Let P be a probability measure on $(\Omega, \mathcal{F})$. A d-dimensional stochastic process starting in 0 with càdlàg paths and independent and stationary increments under P will be called *P-Lévy process* or simply *Lévy process* if there is no ambiguity about the measure involved. $\diamond$

Remark 1.2 This definition of Lévy processes corresponds to the one in Bertoin (1996). Some authors (e.g. Sato (1999), He, Wang and Yan (1992)) drop the requirement of càdlàg paths and define Lévy processes as processes with independent and stationary increments which are stochastically continuous. However, if X is a stochastically continuous process with independent and stationary increments, there exists a càdlàg version of X with the same properties, cf. He, Wang and Yan (1992), Theorem 2.68. Thus our definition is consistent with this definition, as the next two theorems show. Recall that t is called fixed time of discontinuity of a process X if $P[\Delta X_t \neq 0] > 0$. $\diamond$

Theorem 1.3 *A process with càdlàg paths and stationary and independent increments has no fixed times of discontinuity.*

Proof. cf. Jacod and Shiryaev (1987), II.4.3. $\square$

Theorem 1.4 *Let X be a process with stationary and independent increments and càdlàg paths. Then X is stochastically continuous.*

Proof. Suppose X is not stochastically continuous in some $t \geq 0$. Then there exist a sequence $s_n \to t$ and positive ε, δ such that for all $n \in \mathbb{N}$

$$P[|X_t - X_{s_n}| > \varepsilon] \geq \delta.$$

Then since the paths of X are càdlàg and since $s_n \to t$ implies the existence of a monotone subsequence, we have that

$$P[0 > \varepsilon] \geq \delta \text{ or } P[|\Delta X_t| > \varepsilon] \geq \delta,$$

a contradiction to Theorem 1.3. $\square$

If we are given a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in I}$ we call an $\mathbb{F}$-adapted process X with càdlàg paths $(P, \mathbb{F})$-*Lévy process* (or simply $\mathbb{F}$-Lévy process if there is no ambiguity about P) if for $s \leq t$ the random variables $X_t - X_s$ are independent of $\mathcal{F}_s$ under P and if the distribution of $X_t - X_s$ under P depends only on $t-s$ (this corresponds to what Jacod and Shiryaev (1987) call PIIS). Note that if X is a càdlàg process with independent and stationary increments under some measure P (i.e. a P-Lévy process as defined above) and if $\mathbb{F}$ is the filtration generated by X, then X is a $(P, \mathbb{F})$-Lévy process.

We state the following useful properties of Lévy processes for a given stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$.

Theorem 1.5 *Every $\mathbb{F}$-Lévy process is an $\mathbb{F}$-semimartingale.*

Proof. cf. Jacod and Shiryaev (1987), Corollary II.4.19. □

Theorem 1.6 *Let X be an $\mathbb{F}$-Lévy process. Then X is an $\mathbb{F}$-martingale if and only if X is a local $\mathbb{F}$-martingale.*

Proof. cf. He, Wang and Yan (1992), Theorem 11.46. □

Proposition 1.7 *Let X be a càdlàg process and let $\mathbb{F}$ be generated by X. Then X is an $\mathbb{F}$-Lévy process if and only if*

$$E[\exp(iu^{\mathrm{tr}}(X_t - X_s))|\mathcal{F}_s] = E[\exp(iu^{\mathrm{tr}} X_{t-s})]$$

for all $u \in \mathbb{R}^d$ and all $0 \leq s \leq t < \infty$.

Proof. Necessity is obvious. In order to show sufficiency, we take expectations on both sides and get the stationarity of the increments of X. This in turn yields

$$E[\exp(iu^{\mathrm{tr}}(X_t - X_s))|\mathcal{F}_s] = E[\exp(iu^{\mathrm{tr}}(X_t - X_s))],$$

for all $u \in \mathbb{R}^d$, which implies the independence of the increments. □

In the sequel we consider Lévy processes as semimartingales, so we always need some filtration $\mathbb{F}$ on $(\Omega, \mathcal{F})$. The concept is as follows: either we are given a Lévy process L on $(\Omega, \mathcal{F}, P)$, so that a canonical choice of the filtration is the filtration generated by L (in fact we sometimes choose the P-augmentation of $\mathbb{F}^L$); or we are given a filtration $\mathbb{F}$ such that the process L is an $\mathbb{F}$-Lévy process and this is the case when L is considered as a semimartingale.

1.3 Relative Entropy

Relative entropy has lately found its way into mathematical finance, where the concept of maximal expected exponential utility is closely related to minimal relative entropy. Note that despite the fact that relative entropy is not a metric on the space of probability measures, it may be viewed as a measure of distance between two probability measures. Therefore in an incomplete financial market model, where one needs to “choose” a martingale measure from the convex set of martingale measures, the measure with minimal entropy with respect to the real world measure is a natural choice for the pricing measure.

In this section we recall the definition and several properties of relative entropy and define sets of martingale measures over which we shall minimize relative entropy in Part II.

Definition 1.8 Let P, Q be two probability measures on $(\Omega, \mathcal{F})$ and $\mathcal{G}$ some sub-σ-field of $\mathcal{F}$. The *relative entropy of Q with respect to P on $\mathcal{G}$* is defined by

$$I_{\mathcal{G}}(Q|P) = \begin{cases} E_P\left[\frac{dQ}{dP}\Big|_{\mathcal{G}} \log \frac{dQ}{dP}\Big|_{\mathcal{G}}\right] & \text{if } Q|_{\mathcal{G}} \ll P|_{\mathcal{G}} \\ +\infty & \text{otherwise.} \end{cases}$$

For $\mathcal{G} = \mathcal{F}$ we simply write $I_{\mathcal{F}}(Q|P) = I(Q|P)$ and call $I(Q|P)$ the *relative entropy of Q with respect to P.*

If we are given a filtration $\mathbb{F}$ and $Q \stackrel{\text{loc}}{\ll} P$, we define the *entropy process of Q with respect to P* by $I_t(Q|P) = I_{\mathcal{F}_t}(Q|P)$, $t \geq 0$. Note that for a finite time horizon $[0, T]$ and if $\mathcal{F}_T = \mathcal{F}$, $Q \stackrel{\text{loc}}{\ll} P$ is equivalent to $Q \ll P$ and we have $I_{\mathcal{F}_T}(Q|P) = I(Q|P)$. $\diamond$

Lemma 1.9 *Let $Q \ll P$ on some σ-field $\mathcal{G} \subseteq \mathcal{F}$ and let $\mathcal{H} \subseteq \mathcal{G}$ be another σ-field. Then*

$$I_{\mathcal{H}}(Q|P) \leq I_{\mathcal{G}}(Q|P).$$

Proof. The density $Z_{\mathcal{G}} = \frac{dQ}{dP}\Big|_{\mathcal{G}}$ is P-integrable, and without loss of generality we may assume that $I_{\mathcal{G}}(Q|P) < \infty$. Thus $Z_{\mathcal{G}} \log Z_{\mathcal{G}}$ is P-integrable as well, and by Jensen's inequality for conditional expectations we have

$$\begin{aligned} I_{\mathcal{H}}(Q|P) &= E_P\left[Z_{\mathcal{H}} \log Z_{\mathcal{H}}\right] = E_P\left[E_P\left[Z_{\mathcal{G}}|\mathcal{H}\right] \log E_P\left[Z_{\mathcal{G}}|\mathcal{H}\right]\right] \\ &\leq E_P\left[E_P\left[Z_{\mathcal{G}} \log Z_{\mathcal{G}}|\mathcal{H}\right]\right] = E_P\left[Z_{\mathcal{G}} \log Z_{\mathcal{G}}\right] \\ &= I_{\mathcal{G}}(Q|P). \end{aligned}$$

$\square$

Lemma 1.10 *Let $Q, Q' \ll P$ on some σ-field $\mathcal{G} \subseteq \mathcal{F}$. Then for all $0 \leq \lambda \leq 1$*

$$I_{\mathcal{G}}\big((1-\lambda)Q + \lambda Q' \big| P\big) \leq (1-\lambda)\ I_{\mathcal{G}}(Q|P) + \lambda I_{\mathcal{G}}(Q'|P),$$

i.e. relative entropy is a convex functional in the first argument.

Proof. Let $R := (1-\lambda)Q + \lambda Q'$ and Z and Z' be the densities of Q and Q' on $\mathcal{G}$, respectively. Then the density of R on $\mathcal{G}$ is $Z^R = (1-\lambda)Z + \lambda Z'$. Now the function $\varphi(z) = z \log z$ is convex on $[0,\infty)$, so we have

$$\begin{aligned} I_{\mathcal{G}}(R|P) &= E_P\left[\varphi(Z^R)\right] = E_P\left[\varphi((1-\lambda)Z + \lambda Z')\right] \\ &\leq E_P\left[(1-\lambda)\ \varphi(Z) + \lambda\ \varphi(Z')\right] \\ &= (1-\lambda)\ E_P\left[\varphi(Z)\right] + \lambda\ E_P\left[\varphi(Z')\right] \end{aligned}$$

and hence the claim. □

Lemma 1.11 *Let $Q \overset{\text{loc}}{\ll} P$ with density process Z and finite-valued entropy process. Then $Z \log Z$ is a P-submartingale and $\log Z$ is a Q-submartingale.*

Proof. Let $0 \leq s \leq t$. Note that $Z_t \log Z_t$ is P-integrable due to the finiteness of $I_t(Q|P)$. So by Jensen's inequality and the fact that Z is a P-martingale we have

$$E_P[Z_t \log Z_t | \mathcal{F}_s] \geq E_P[Z_t|\mathcal{F}_s]\ \log E_P[Z_t|\mathcal{F}_s] = Z_s \log Z_s$$

P-a.s., hence Q-a.s. Now $E_Q[\log Z_t|\mathcal{F}_s] = \frac{1}{Z_s} E_P[Z_t \log Z_t|\mathcal{F}_s]$, so it is immediate that $\log Z$ is a Q-submartingale. □

Note that since $Z \log Z$ is a submartingale the entropy process of some $Q \overset{\text{loc}}{\ll} P$ is increasing if it is finite-valued, so that in the case of a finite time horizon $[0,T]$ the finiteness of the entropy process is equivalent to the finiteness of the relative entropy.

Definition 1.12 Let $Q^n \overset{\text{loc}}{\ll} P$ for $n \in \mathbb{N}$ and $Q \overset{\text{loc}}{\ll} P$ with $I_t(Q|P) < \infty$ for all $t \geq 0$. We say that Q^n *converges in entropy to* Q if $\lim_{n\to\infty} I_t(Q^n|P) = I_t(Q|P)$ for all $t \geq 0$. ◇

Remark 1.13 This definition is not coherent with what is usually found in the literature. In fact, some authors say that Q^n converges to Q in entropy if $I(Q^n|Q) \overset{n\to\infty}{\longrightarrow} 0$. Also note that in our definition convergence in entropy does not necessarily imply uniqueness of the limiting measure neither does it induce a topology on $\{Q \mid Q \overset{\text{loc}}{\ll} P\}$. ◇

Definition 1.14 Let X be a d-dimensional semimartingale and let U be a $d \times d$-matrix.

a) We define the following sets of local martingale measures for UX.

$$\begin{aligned} \mathcal{Q}_a^U(X) &:= \left\{ Q \overset{\text{loc}}{\ll} P \;\middle|\; UX \text{ is a local } Q\text{-martingale} \right\} \\ \mathcal{Q}_e^U(X) &:= \left\{ Q \in \mathcal{Q}_a^U(X) \;\middle|\; Q \overset{\text{loc}}{\sim} P \right\} \\ \mathcal{Q}_f^U(X) &:= \left\{ Q \in \mathcal{Q}_a^U(X) \mid I_t(Q|P) < \infty \text{ for } t \geq 0 \right\}. \end{aligned}$$

If X is a P-Lévy process, we define

$$\mathcal{Q}_\ell^U(X) \quad := \quad \left\{Q \in \mathcal{Q}_a^U(X) \mid X \text{ is a } Q\text{-Lévy process}\right\}.$$

Note that $Q \in \mathcal{Q}_\ell^U(X)$ means that the transformed process UX is a local Q-martingale whereas we require the *original* process X to be a Lévy process. This will be important when U is a projection matrix or, more generally, when U is not regular. We call $Q \in \mathcal{Q}_\ell^U(X)$ a *Lévy martingale measure for* UX or simply Lévy martingale measure if there is no ambiguity about X and U. If U is the identity matrix, we omit the dependence on U and simply write $\mathcal{Q}_x = \mathcal{Q}_x^U$, $x \in \{a, e, f, \ell\}$.

b) The *entropy-minimizing local martingale measure* $Q^E(X)$ is defined as the measure which minimizes the entropy process pointwise over all $Q \in \mathcal{Q}_a(X)$, i.e.

$$I_t(Q^E(X)|P) \leq I_t(Q|P) \text{ for all } Q \in \mathcal{Q}_a(X), t \geq 0.$$

Provided we have the existence of $Q^E(X)$, convexity of the relative entropy yields uniqueness of $Q^E(X)$ on $\mathcal{F}_t$ for all $t \geq 0$, and thus uniqueness of $Q^E(X)$ on $\mathcal{F}_\infty$.

c) We define the *entropy-minimizing Lévy martingale measure for* UL as the measure which minimizes the entropy process pointwise over all $Q \in \mathcal{Q}_\ell^U$, i.e.

$$I_t(Q_\ell^E|P) \leq I_t(Q|P) \text{ for all } Q \in \mathcal{Q}_\ell^U, t \geq 0.$$

Provided we have the existence of Q_ℓ^E, convexity of the relative entropy implies again uniqueness of Q_ℓ^E. ◇

In Frittelli (2000), Theorem 2.2, it is shown that if there exists an equivalent martingale measure with finite relative entropy, then $Q^E(X)$, if it exists, is equivalent to P on $\mathcal{F}_\infty$. We will see that in the case of an infinite time horizon, the relative entropy of a *locally* equivalent martingale measure for a P-Lévy process L is in general infinite, but we have the following extension of these results.

Lemma 1.15 *If $Q^E(X)$ exists and $\mathcal{Q}_e(X) \cap \mathcal{Q}_f(X) \neq \emptyset$, then $Q^E(X) \in \mathcal{Q}_e(X) \cap \mathcal{Q}_f(X)$.*

Proof. Let $t \geq 0$, then $Q^E(X)\big|_{\mathcal{F}_t}$ is the martingale measure which minimizes relative entropy on $\mathcal{F}_t$. Furthermore $\mathcal{Q}_e(X) \cap \mathcal{Q}_f(X) \neq \emptyset$ implies that there exists $\bar{Q} \stackrel{\text{loc}}{\sim} P$ with $I_t(\bar{Q}|P) < \infty$ and $\bar{Q}\big|_{\mathcal{F}_t} \sim P|_{\mathcal{F}_t}$ on $\mathcal{F}_t$. Then by Frittelli (2000), Theorem 2.2, $Q^E(X)\big|_{\mathcal{F}_t} \sim P|_{\mathcal{F}_t}$ on $\mathcal{F}_t$, hence $Q^E(X) \stackrel{\text{loc}}{\sim} P$. □

1.4 Characteristics of Semimartingales

In this section we recall the notion of characteristics of semimartingales as presented in Jacod and Shiryaev (1987). Characteristics are a generalization of the characteristic triplet of an infinitely divisible distribution, just like the concept of semimartingales may be viewed as a generalization of Lévy processes. However note that in general the characteristics of a semimartingale do not uniquely describe the distribution of a semimartingale. Roughly speaking, the characteristics describe drift, volatility and jumps of a semimartingale. We first give a short introduction on random measures and the concept of stochastic integral with respect to random measures before we arrive at characteristics of semimartingales, their properties and several examples. Recall that all semimartingales are taken to have càdlàg paths.

Random Measures

Random measures and their compensators are a useful tool to capture the behaviour of the jumps of a semimartingale. We begin with the following general definition of random measures before turning our attention to what is needed for characteristics of semimartingales, namely random measures associated with the jumps of a semimartingale.

Definition 1.16 a) A *random measure* on $\mathbb{R}_+ \times \mathbb{R}^d$ is a family

$$\mu = \{\mu(\omega; dt, dx) \colon \omega \in \Omega\}$$

of measures on $(\mathbb{R}_+ \times \mathbb{R}^d, \mathcal{B}_+ \otimes \mathcal{B}^d)$ satisfying $\mu(\omega; \{0\} \times \mathbb{R}^d) = 0$ for all $\omega \in \Omega$.

b) Let μ be a random measure and let W be an optional function. We introduce the integral process $W * \mu$ by

$$W * \mu_t(\omega) = \begin{cases} \int_{[0,t]\times\mathbb{R}^d} W(\omega, s, x)\ \mu(\omega, ds, dx) & \text{if the integral converges} \\ +\infty & \text{otherwise.} \end{cases}$$

c) A random measure μ is called *optional* (respectively *predictable*) if the process $W * \mu$ is optional (respectively predictable) for every optional (respectively predictable) function W. $\diamond$

We now introduce the random measures which are the most important in our context, namely random measures associated with the jumps of a semimartingale, and their compensators.

Definition 1.17 Let X be a semimartingale. The random measure μ^X associated with the jumps or *jump measure* of X is defined by (cf. Jacod and Shiryaev (1987), Proposition II.1.16.)

$$\mu^X(dt, dx) = \sum_{s>0} \mathbb{1}_{\{\Delta X_s \neq 0\}}\ \delta_{\{(s,\Delta X_s)\}}(dt, dx).$$

Note that μ^X is integer-valued and optional. $\diamond$

Remark 1.18 Let μ^X be the jump measure of X. For any nonnegative optional function W we have

$$W * \mu_t^X = \sum_{0<s\leq t} W(s, \Delta X_s) \; 1\!\!1_{\{\Delta X_s \neq 0\}}.$$

(cf. Jacod and Shiryaev (1987), II.1.15.) ◇

Theorem and Definition 1.19 Let μ^X be the jump measure of X. The *predictable P-compensator of* μ^X, denoted by ν^P (or simply ν if there is no ambiguity about the measure involved), is the predictable random measure which satisfies one of the two following equivalent conditions:

(i) $E_P[W * \nu^P_\infty] = E_P[W * \mu^X_\infty]$ for every nonnegative predictable function W.

(ii) For every predictable function W such that $|W| * \mu^X$ is finite-valued and locally P-integrable (which is equivalent to $|W| * \nu^P$ being finite-valued and locally P-integrable), $W * \mu^X - W * \nu^P$ is a local P-martingale.

Note that ν^P is unique up to a P-null set.

Proof. cf. Jacod and Shiryaev (1987), Theorem II.1.8. □

We now define the stochastic integral with respect to a compensated random measure. As we will see this coincides with the definition of the stochastic integral with respect to a purely discontinuous local martingale. To that end, let X be a semimartingale without fixed times of discontinuity, and let μ^X be the random measure associated with the jumps of X and ν^P its compensator. Furthermore, let W be a predictable function and define the stochastic process $\widetilde{W}$ by $\widetilde{W}_s := W(s, \Delta X_s) 1\!\!1_{\{\Delta X_s \neq 0\}}$, $s \geq 0$.

Definition 1.20 a) We say that W is *integrable* with respect to $\mu^X - \nu^P$ if the increasing process $\left(\sum_{s\leq \cdot} \widetilde{W}_s^2\right)^{1/2}$ is locally P-integrable.

b) If W is integrable with respect to $\mu^X - \nu^P$, we call *stochastic integral* of W with respect to $\mu^X - \nu^P$ any purely discontinuous local martingale M such that the processes ΔM and $\widetilde{W}$ are indistinguishable. The stochastic integral is then denoted by $W * (\mu^X - \nu^P)$. ◇

Theorem 1.21 *a) If the increasing process $|W| * \mu^X$ (or, equivalently, $|W| * \nu^P$) is locally P-integrable, then W is integrable with respect to $\mu^X - \nu^P$ and*

$$W * (\mu^X - \nu^P) = W * \mu^X - W * \nu^P.$$

b) *Let H be a locally bounded predictable process and let W be integrable with respect to $\mu^X - \nu^P$. Then HW is integrable with respect to $\mu^X - \nu^P$ and*

$$(HW) * (\mu^X - \nu^P) = \int H \; d(W * (\mu^X - \nu^P)).$$

Proof. cf. Jacod and Shiryaev (1987), Propositions II.1.28 and II.1.30. □

Theorem 1.22 *Suppose that X is such that $\nu^P(\{t\} \times \mathbb{R}^d) = 0$ P-a.s. for each $t \geq 0$. (We shall see later that this holds if X is a Lévy process.) Then:*

a) *W is integrable with respect to $\mu^X - \nu^P$ and $W * (\mu^X - \nu^P)$ is a locally square-integrable P-martingale if and only if the increasing process $W^2 * \nu^P$ is locally integrable, in which case $\langle W * (\mu^X - \nu^P) \rangle = W^2 * \nu^P$.*

b) *W is integrable with respect to $\mu^X - \nu^P$ and $W * (\mu^X - \nu^P)$ is of (locally) integrable variation if and only if the increasing process $|W| * \nu^P$ is (locally) integrable.*

c) *Assume $\widetilde{W} \geq -1$ identically. Then W is integrable with respect to $\mu^X - \nu^P$ if and only if the increasing process*

$$\left(1 - \sqrt{1 + W}\right)^2 * \nu^P$$

is locally P-integrable.

Proof. cf. Jacod and Shiryaev (1987), Theorem II.1.33; note that our assumption on X implies that $a \equiv 0$ and $\hat{W} \equiv 0$ there. □

Characteristics of Semimartingales

We now introduce characteristics of semimartingales, a concept heavily used in Part II. The idea is to associate to a semimartingale a triplet of predictable processes which describe drift, volatility and jumps, in analogy to the concept of characteristic triplet of infinitely divisible distributions, which in turn describes drift, volatility and jumps of the associated Lévy process.

Let X be a semimartingale and h a truncation function and define the process $X(h)$ by

$$X(h)_t = X_t - \sum_{s \leq t} (\Delta X_s - h(\Delta X_s)).$$

Note that $\sum_{s \leq t} (\Delta X_s - h(\Delta X_s)) = (x - h(x)) * \mu_t^X$, and since $\Delta X_s - h(\Delta X_s) \neq 0$ only for finitely many s, this sum converges. Furthermore $\Delta X(h) = h(\Delta X)$ is bounded, so $X(h)$ is a special semimartingale with canonical decomposition

$$X(h) = X_0 + M(h) + B(h), \tag{1.1}$$

where $M(h)$ is a local P-martingale and $B(h)$ is a predictable process with finite variation. Let μ^X be the random measure associated with the jumps of X.

Definition 1.23 The triplet (B, C, ν) with

$$\begin{array}{rcll} B &=& B(h) & \text{from the canonical decomposition (1.1)} \\ C &=& (\langle X^{i,c}, X^{j,c} \rangle)_{1 \le i,j \le d} & \\ \nu &=& \nu^P, & \text{the } (P\text{-}) \text{ compensator of } \mu^X \end{array}$$

is called the triplet of *P-characteristics* of X relative to the truncation function h or simply characteristics if there is no ambiguity about the measure and the truncation function involved. Sometimes B is called the *first*, C the *second*, and ν the *third characteristic* of X.

Concerning limit theorems for stochastic processes it is necessary to define the *modified second characteristic* of X relative to h as the predictable process $\tilde{C}$ with

$$\tilde{C}^{ij} = \langle M(h)^i, M(h)^j \rangle,$$

where $M(h)$ is the martingale part in the decomposition (1.1). ◇

Remark 1.24 Obviously only the first and the modified second characteristic depend on the choice of the truncation function. In the sequel we fix one truncation function and sometimes do not mention the dependence of the characteristics on this truncation function. If X is already a special semimartingale, one can directly define the characteristics using the "truncation function" $h(x) = x$. ◇

Proposition 1.25 *Let X be a semimartingale with characteristics B, C, ν relative to some truncation function h. Then the modified second characteristic $\tilde{C}$ is given by*

$$\tilde{C}_t^{ij} = C_t^{ij} + (h^i h^j) * \nu_t - \sum_{s \le t} \Delta B_s^i \Delta B_s^j$$

up to an evanescent set.

Proof. cf. Jacod and Shiryaev (1987), Proposition II.2.17. □

Theorem 1.26 *Let X be a semimartingale. Then there exists a version of the characteristics of X which is of the form*

$$\begin{array}{rcl} B^i &=& \int b^i \, dA, \\ C^{ij} &=& \int c^{ij} \, dA, \\ \nu(\omega; dt, dx) &=& dA_t(\omega) \; K_{\omega,t}(dx), \end{array}$$

where A is a real-valued predictable, increasing and locally integrable process, $b = (b^i)_{1 \le i \le d}$ is an $\mathbb{R}^d$-valued predictable process, $c = (c^{ij})_{1 \le i,j \le d}$ is a predictable process with values in the set *of all symmetric nonnegative definite $d \times d$-matrices, and $K_{\omega,t}(dx)$ is a transition kernel from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}^d, \mathcal{B}^d)$ which satisfies for each t*

$$K_{\omega,t}(\{0\}) = 0 \text{ and } \int_{\mathbb{R}^d} (|x|^2 \wedge 1)\ K_{\omega,t}(dx) \le 1.$$

Proof. cf. Jacod and Shiryaev (1987), Proposition II.2.9. □

Theorem 1.27 *Let X be a semimartingale. Then*

a) X is a process with independent increments if and only if its characteristics are deterministic.

b) X is a Lévy process if and only if its characteristics are deterministic and linear in time, i.e.

$$\begin{aligned} B_t &= bt, \\ C_t &= ct, \\ \nu(dt, dx) &= dtK(dx), \end{aligned}$$

where $b \in \mathbb{R}^d$, c is a symmetric nonnegative definite $d \times d$-matrix, and K is a σ-finite measure on $(\mathbb{R}^d, \mathcal{B}^d)$ with $\int_{\mathbb{R}^d}(1 \wedge |x|^2)\ K(dx) < \infty$ and $K(\{0\}) = 0$.

Proof. cf. Jacod and Shiryaev (1987), Theorem II.4.15 and Corollary II.4.19. The assertion that K is σ-finite is not explicitly stated there but follows from the fact that a measure μ is σ-finite if and only if there exists $\varphi \in L^2(\mu)$ with $\varphi > 0$ μ-a.e. If we take $\varphi(x) = (1 \wedge |x|^2)^{1/2}$, then $\varphi > 0$ K-a.e. since $K(\{0\}) = 0$, and $\varphi \in L^2(K)$ since $\int_{\mathbb{R}^d}(1 \wedge |x|^2)\ K(dx) < \infty$. □

Remark 1.28 Theorem 1.27 b) is a special case of Theorem 1.26. It states that in the Lévy case one can choose $A_t = t$, the processes b and c constant and K independent of ω and t, thus making it a measure. If X is a Lévy process, we also call (b, c, K) the *Lévy characteristics* of X. Note that the triplet (b, c, K) coincides with the characteristic triplet from the Lévy-Khinchine representation of the infinitely divisible distribution of X_1, cf. Theorems A.3 and A.5 in the appendix.

Furthermore, Theorem 1.27 b) shows that for a Lévy process X with compensator ν^P of the jump measure μ^X, we have $\nu^P(\{t\} \times \mathbb{R}^d) = 0$ P-a.s. for all $t \ge 0$. Hence Lévy processes satisfy the assumption of Theorem 1.22. ◇

Theorem 1.29 *Let X be a semimartingale. If X has independent increments, then the (deterministic) characteristics of X uniquely determine the distribution of X.*

Proof. cf. Jacod and Shiryaev (1987), Theorem II.4.25. □

Proposition 1.30 *Let X be a semimartingale with characteristics (B, C, ν). Then X is a special semimartingale if and only if $(|x|^2 \wedge |x|) * \nu$ is locally integrable.*

Proof. cf. Jacod and Shiryaev (1987), Proposition II.2.29. □

Theorem 1.31 *Let X be a semimartingale with characteristics (B, C, ν) relative to a truncation function h. Then h is integrable with respect to $\mu^X - \nu$ and the following representation holds:*

$$X = X_0 + X^c + h * (\mu^X - \nu) + (x - h(x)) * \mu^X + B.$$

Proof. cf. Jacod and Shiryaev (1987), Theorem II.2.34. □

Corollary 1.32 *Let X be a special semimartingale with characteristics (B, C, ν) relative to the "truncation function" $h(x) = x$. Then x is integrable with respect to $\mu^X - \nu$ and we have*

$$X = X_0 + X^c + x * (\mu^X - \nu) + B.$$

Proof. cf. Jacod and Shiryaev (1987), Corollary II.2.38. □

Corollary 1.33 *a) Let h be a fixed truncation function. A semimartingale X with characteristics (B, C, ν) relative to h is a local martingale if and only if $B + (x - h(x)) * \mu^X$ is a local martingale.*

b) A special semimartingale X with characteristics (B, C, ν) relative to the "truncation function" $h(x) = x$ is a local martingale if and only if $B \equiv 0$.

Proof. This follows immediately from Theorem 1.31 and Corollary 1.32. □

The following theorem is concerned with the characteristics of a linear transformation of a semimartingale. For a $d \times d$-matrix U and $A \subseteq \mathbb{R}^d$ we define $U^{-1}(A) = \{x \in \mathbb{R}^d \mid Ux \in A\}$.

Theorem 1.34 *Let X be a d-dimensional semimartingale with characteristics (B, C, ν) relative to a truncation function h and let U be a $d \times d$-matrix. Then the semimartingale $\tilde{X}$, given by $\tilde{X}_t = UX_t$, admits the following characteristics relative to h:*

$$\begin{aligned} B_t^{\tilde{X}} &= UB_t - U_h * \nu_t, \\ C_t^{\tilde{X}} &= UC_tU^{\mathrm{tr}}, \\ \nu^{\tilde{X}}(A_1 \times A_2) &= \nu\left(A_1 \times U^{-1}(A_2 \setminus \{0\})\right), \quad A_1 \in \mathcal{B}(\mathbb{R}_+), A_2 \in \mathcal{B}(\mathbb{R}^d), \end{aligned}$$

where $U_h \colon \mathbb{R}^d \to \mathbb{R}^d$ with $U_h(x) = Uh(x) - h(Ux)$.

Proof. It is immediate that the linear transformation of a semimartingale is again a semimartingale. Concerning the characteristics of $\tilde{X}$ we follow the definition of characteristics. $B^{\tilde{X}}$ is given by the canonical decomposition of the special semimartingale $\tilde{X}(h)$, where

$$\begin{aligned}
\tilde{X}(h)_t &= \tilde{X}_t - (x - h(x)) * \mu_t^{\tilde{X}} \\
&= UX_t - \sum_{s \le t} (\Delta(UX)_s - h(\Delta(UX)_s)) \\
&= U\left(X_t - \sum_{s \le t} (\Delta X_s - h(\Delta X_s))\right) - \sum_{s \le t} (Uh(\Delta X_s) - h(U\Delta X_s)) \\
&= UX(h)_t - U_h * \mu_t^X.
\end{aligned}$$

Now the increasing process $|U_h| * \mu^X$ is locally integrable. Indeed $Uh(x) - h(Ux) \neq 0$ implies $\varepsilon \le |x| \le R$ for some $0 < \varepsilon, R < \infty$. Let $\varepsilon' > 0$ be such that $h(x) = x$ for $|x| < \varepsilon'$. Then there exists $\varepsilon'' > 0$ such that $|Ux| < \varepsilon'$ for $|x| < \varepsilon''$ since U is linear on $\mathbb{R}^d$ and thus uniformly continuous. So if we choose $\varepsilon = \varepsilon' \wedge \varepsilon''$ we have $Uh(x) - h(Ux) = 0$ for $|x| < \varepsilon$. With similar arguments and the fact that $h(x) = 0$ for $|x| > R'$ we get $Uh(x) - h(Ux) = 0$ for $|x| > R$ for some $R < \infty$. Thus $|Uh(x) - h(Ux)| * \mu_t^X$ is a finite sum over bounded random variables, hence $|Uh(x) - h(Ux)| * \mu^X$ is locally integrable. With this we can decompose $(Uh(x) - h(Ux)) * \mu^X = U_h * \mu^X$ as

$$U_h * \mu^X = U_h * (\mu^X - \nu) + U_h * \nu,$$

so the canonical decomposition of the special semimartingale $\tilde{X}(h) = \tilde{M}(h) + \tilde{B}(h)$ is given by

$$\begin{aligned}
\tilde{M}(h)_t &= UM(h)_t - U_h * (\mu^X - \nu)_t, \\
\tilde{B}(h)_t &= UB(h)_t - U_h * \nu_t,
\end{aligned}$$

which yields the claimed form of $B^{\tilde{X}}$. The special form of $C^{\tilde{X}}$ is immediate by the uniqueness of the continuous martingale part of a semimartingale and

$$C^{\tilde{X}} = \langle \tilde{X}^c \rangle = \langle UX^c \rangle = UCU^{\mathrm{tr}}.$$

Concerning the third characteristic of $\tilde{X}$, we see that for any nonnegative predictable function $\tilde{W} \colon \Omega \times \mathbb{R}_+ \times \mathbb{R}^d$ we have

$$\begin{aligned}
\tilde{W} * \mu_t^{\tilde{X}} &= \sum_{s \le t} \tilde{W}(s, \Delta\tilde{X}_s)\ \mathbb{1}_{\{|\Delta\tilde{X}_s| \neq 0\}} \\
&= \sum_{s \le t} \tilde{W}(s, \Delta UX_s)\ \mathbb{1}_{\{|\Delta UX_s| \neq 0\}} \\
&= \sum_{s \le t} \tilde{W}(s, U\Delta X_s)\ \mathbb{1}_{\{|U\Delta X_s| \neq 0\}} \mathbb{1}_{\{|\Delta X_s| \neq 0\}} \\
&= W * \mu_t^X,
\end{aligned}$$

where $W(s,x) := \tilde{W}(s,Ux)\mathbb{1}_{\{|Ux|\neq 0\}}$. Now let $\nu^{\tilde{X}}$ be the compensator of $\mu^{\tilde{X}}$, then

$$E[\tilde{W} * \nu^{\tilde{X}}_{\infty}] = E[\tilde{W} * \mu^{\tilde{X}}_{\infty}] = E[W * \mu^{X}_{\infty}] = E[W * \nu_{\infty}],$$

since ν is the compensator of μ^X. Now by Jacod and Shiryaev (1987), Theorem I.2.2, the predictable σ-field $\mathcal{P}$ on $\Omega \times \mathbb{R}_+$ is generated by the family $\{A \times (t,u] \mid t < u, A \in \mathcal{F}_t\}$, which implies that $\widetilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ is generated by $\{A \times (t,u] \times A_2 \mid t < u, A \in \mathcal{F}_t, A_2 \in \mathcal{B}(\mathbb{R}^d)\}$. So we fix $t < u$ and we define the predictable function $\tilde{W}$ by $\tilde{W}(\omega; s, x) = \mathbb{1}_A(\omega)\mathbb{1}_{A_1}(s)\mathbb{1}_{A_2}(x)$ for $A \in \mathcal{F}_t$, $A_1 = (t,u]$ and $A_2 \in \mathcal{B}(\mathbb{R}^d)$. Then for W as defined above we obtain

$$\begin{aligned}
W(\omega; s, x) &= \tilde{W}(\omega, s, Ux)\ \mathbb{1}_{\{|Ux|\neq 0\}}(x) \\
&= \mathbb{1}_A(\omega)\ \mathbb{1}_{A_1}(s)\ \mathbb{1}_{A_2}(Ux)\ \mathbb{1}_{\{|Ux|\neq 0\}}(x) \\
&= \mathbb{1}_A(\omega)\ \mathbb{1}_{A_1}(s)\ \mathbb{1}_{A_2 \setminus \{0\}}(Ux) \\
&= \mathbb{1}_A(\omega)\ \mathbb{1}_{A_1}(s)\ \mathbb{1}_{U^{-1}(A_2 \setminus \{0\})}(x).
\end{aligned}$$

Altogether we get for A, A_1 and A_2 as above

$$E\left[\mathbb{1}_A\ \nu^{\tilde{X}}(A_1 \times A_2)\right] = E\left[\tilde{W} * \nu^{\tilde{X}}_{\infty}\right] = E\left[W * \nu_{\infty}\right] = E\left[\mathbb{1}_A\ \nu(A_1 \times U^{-1}(A_2 \setminus \{0\}))\right],$$

and since $\mathcal{F} = \mathcal{F}_{\infty}$ this yields the assertion. □

Theorem 1.34 deserves some comments. With its help we can, for example, determine the characteristics of projections of a d-dimensional semimartingale X on the coordinate axes. It is worthwhile to note that these are not simply the projections of the characteristics of X. In fact we get an additional drift term, and the compensator of the jump measure does not take into account the jumps which lie in the kernel of U, i.e. the jumps which are orthogonal to the subspace of $\mathbb{R}^d$ on which the projection takes place.

For a projection on $\mathbb{R}^n$ one could also take U to be an $n \times d$-matrix, however then UX is an n-dimensional semimartingale and one has to consider an additional truncation function on $\mathbb{R}^n$, so for notational convenience we think of projections as mappings from $\mathbb{R}^d$ to $\mathbb{R}^d$.

Compare Theorem 1.34 to Sato (1999), Proposition 11.10, where an analogue for Lévy processes is stated which we state as a corollary.

Corollary 1.35 *Let L be a Lévy process with Lévy characteristics (b, c, K) and U a $d \times d$-matrix. Then UL is a Lévy process with Lévy characteristics $\tilde{b}, \tilde{c}, \tilde{K}$, where*

$$\begin{aligned}
\tilde{b} &= Ub - \int_{\mathbb{R}^d} U_h(x)\ K(dx), \\
\tilde{c} &= UcU^{\mathrm{tr}}, \\
\tilde{K}(A) &= K\left(U^{-1}(A \setminus \{0\})\right), \quad A \in \mathcal{B}(\mathbb{R}^d).
\end{aligned}$$

Proof. By Theorem 1.27 we know that the characteristics of L are deterministic and linear in time, and $\nu(ds, dx) = dsK(dx)$. So by Theorem 1.34 the characteristics of UL are given by

$$\begin{aligned}
\tilde{B}_t &= \left(Ub - \int_{\mathbb{R}^d} U_h(x)\ K(dx) \right) t, \\
\tilde{C}_t &= UcU^{\mathrm{tr}}\ t, \\
\tilde{\nu}([0,t] \times A) &= K\left(U^{-1}(A \setminus \{0\})\right) t,
\end{aligned}$$

which yields the claim. □

Theorem 1.36 *Let X be a semimartingale with characteristics (B, C, ν). If $C_t = ct$ for some deterministic matrix c, then $X^c \stackrel{d}{=} A\ W$, where W is a d-dimensional standard Brownian motion and A is a matrix such that $AA^{\mathrm{tr}} = c$.*

Proof. By definition of the characteristics of X, the continuous martingale part X^c is a continuous local martingale with $\langle X^c \rangle_t = ct$. Let λ^i be the (nonnegative) eigenvalues of c. Since c is symmetric and nonnegative definite, there exists a diagonal matrix $\tilde{c}$ with $\tilde{c}^{ii} = \lambda^i$ and an orthogonal matrix S such that $c = S\tilde{c}S^{\mathrm{tr}}$. Define the matrix $\sqrt{\tilde{c}}$ by $\left(\sqrt{\tilde{c}}\right)^{ij} = \sqrt{\tilde{c}^{ij}}$, then

$$c = S\tilde{c}S^{\mathrm{tr}} = S\sqrt{\tilde{c}}\left(S\sqrt{\tilde{c}}\right)^{\mathrm{tr}} =: AA^{\mathrm{tr}}.$$

Now the characteristics of X^c are $(0, ct, 0)$ and deterministic, so that X^c is a Lévy process. On the other hand the quadratic variation of $A\ W$ is

$$\langle A\ W \rangle_t = A\langle W \rangle_t A^{\mathrm{tr}} = AA^{\mathrm{tr}}t = ct.$$

Since $A\ W$ is a continuous local martingale (and thus a special semimartingale) and a Lévy process, its characteristics are $(0, ct, 0)$, so that the characteristics of X^c and $A\ W$ coincide. Hence $X^c \stackrel{d}{=} A\ W$ by Theorem 1.29. □

1.5 Convergence of Semimartingales

In this section we recall some known results concerning tightness and convergence of sequences of semimartingales. As we are concerned with convergence in distribution we suppose that for a sequence $(X^n)_{n\in\mathbb{N}}$ of processes, X^n is defined on a stochastic basis $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, P^n)$. If furthermore a process X is defined on some $(\Omega, \mathcal{F}, \mathbb{F}, P)$ we denote convergence in distribution of X^n to X, i.e. $\mathcal{L}(X^n|P^n) \xrightarrow{w} \mathcal{L}(X|P)$, by $X^n \xrightarrow{\mathcal{L}} X$ if there is no ambiguity about P^n and P.

Continuous Mapping and Skorokhod Embedding

We first recall two well-known results concerning weak convergence, namely the continuous mapping theorem and the Skorokhod embedding theorem.

Theorem 1.37 (Continuous Mapping Theorem) *Let (S,d) and (S',d') be two metric spaces, endowed with the Borel-σ-algebras $\mathcal{B}_S$ and $\mathcal{B}_{S'}$, respectively, and let $\mu, (\mu^n)_{n\in\mathbb{N}}$ be probability measures on $(S,\mathcal{B}_S)$. Let furthermore $\varphi^n, \varphi\colon S \to S'$ be a sequence of measurable functions and denote by D the set of all $x \in S$ such that there exists a sequence $(x^n)_{n\in\mathbb{N}}$ with $x^n \to x$ but $\varphi^n(x^n) \not\to \varphi(x)$. If S' is separable, then $D \in \mathcal{B}_{S'}$, and in this case the assumptions $\mu^n \xrightarrow{w} \mu$ and $\mu(D) = 0$ imply $\mu^n(\varphi^n)^{-1} \xrightarrow{w} \mu\varphi^{-1}$.*

Proof. cf. Billingsley (1968), Theorem 5.5. □

Remark 1.38 a) If $\varphi^n = \varphi$ for all n, Theorem 1.37 reduces to the classical continuous mapping theorem: If φ is μ-a.e. continuous, then $\mu^n \xrightarrow{w} \mu$ implies $\mu^n\varphi^{-1} \xrightarrow{w} \mu\varphi^{-1}$.

b) If μ^n and μ are the distributions of random variables X^n and X under some measures P^n and P, then the continuous mapping theorem carries over to convergence in distribution in the following sense. If $X^n \xrightarrow{\mathcal{L}} X$ and if $P[X \in D] = 0$, then $\varphi^n(X^n) \xrightarrow{\mathcal{L}} \varphi(X)$. If $\varphi^n = \varphi$ for all $n \in \mathbb{N}$ the condition $P[X \in D] = 0$ means that φ is PX^{-1}-a.e. continuous. ◇

Theorem 1.39 (Skorokhod Embedding Theorem) *Let (S,d) be a separable metric space endowed with the Borel-σ-algebra $\mathcal{B}_S$ and let $(\mu^n)_{n\in\mathbb{N}}, \mu$ be probability measures on $(S,\mathcal{B}_S)$ with $\mu^n \xrightarrow{w} \mu$. Then there exist a probability space $(\Omega,\mathcal{F},P)$ and S-valued random variables X and X^n, all defined on $(\Omega,\mathcal{F},P)$ with distributions μ and μ^n, respectively, and such that $X^n \to X$ P-a.s.*

Proof. cf. Ethier and Kurtz (1986), Theorem 3.1.8. □

There are some aspects where one needs to be careful when using the Skorokhod embedding theorem. For example if $(X^n, Y^n) \xrightarrow{\mathcal{L}} (X,Y)$, then there are random variables $(\hat{X}^n, \hat{Y}^n)$ and $(\hat{X}, \hat{Y})$ on a common probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ with the same distributions as (X^n, Y^n) and (X,Y) under P^n and P, but how far do certain dependences which are not immediately manifest in the common distribution of X^n and Y^n carry over to $\hat{X}^n$ and Y^n? We have the following results.

Lemma 1.40 *Let X and Y be random variables on some probability space $(\Omega,\mathcal{F},P)$ and with values in some separable metric space (S,d). Let $\hat{X}$ and $\hat{Y}$ be defined on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ such that $\mathcal{L}((X,Y)|P) = \mathcal{L}((\hat{X},\hat{Y})|\hat{P})$. If $Y = f(X)$ P-a.s. for some measurable function $f\colon S \to S$, then $\hat{Y} = f(\hat{X})$ $\hat{P}$-a.s.*

Proof. Since (S,d) is separable, $d(X,Y)$ is a random variable, and we have

$$\begin{aligned}&\hat{P}[\hat{Y}=f(\hat{X})]=\hat{P}[d(\hat{Y},f(\hat{X}))=0]\\ =\ &\hat{P}[(\hat{X},\hat{Y})\in g^{-1}(\{0\})]=P[(X,Y)\in g^{-1}(\{0\})]=P[Y=f(X)]=1,\end{aligned}$$

for $g\colon S\times S\to\mathbb{R}$, $g(x,y)=d(x,f(y))$. □

Lemma 1.41 *Let X, Y and Z be random variables on $(\Omega,\mathcal{F},P)$ with values in some separable metric space (S,d). Let $\mathcal{G}=\sigma(Z)$ and suppose $Y=E_P[X|\mathcal{G}]$ P-a.s. If $\hat{X}$, $\hat{Y}$ and $\hat{Z}$ are random variables on $(\hat{\Omega},\hat{\mathcal{F}},\hat{P})$ with $\mathcal{L}((X,Y,Z)|P)=\mathcal{L}((\hat{X},\hat{Y},\hat{Z})|\hat{P})$, then $\hat{Y}=E_{\hat{P}}[\hat{X}|\hat{\mathcal{G}}]$ $\hat{P}$-a.s., where $\hat{\mathcal{G}}=\sigma(\hat{Z})$.*

Proof. Since $Y=E_P[X|\mathcal{G}]$ and $\mathcal{G}$ is generated by Z, we have that $Y=f(Z)$ for some measurable function f. So by Lemma 1.40 we have $\hat{Y}=f(\hat{Z})$, so that $\hat{Y}$ is $\hat{\mathcal{G}}$-measurable. It remains to show that for all $\hat{A}\in\hat{\mathcal{G}}$ we have $E_{\hat{P}}[\hat{Y}\mathbb{1}_{\hat{A}}]=E_{\hat{P}}[\hat{X}\mathbb{1}_{\hat{A}}]$. Now since $\hat{\mathcal{G}}$ is generated by $\hat{Z}$, we have that $\hat{A}=\hat{Z}^{-1}(B)$ for some $B\in\mathcal{B}_S$. Therefore we have with $g_B\colon S\times S\to S$, $g_B(y,z)=y\mathbb{1}_B(z)$, and $A=Z^{-1}(B)\in\mathcal{G}$ that

$$\begin{aligned}&E_{\hat{P}}[\hat{Y}\mathbb{1}_{\hat{A}}]=E_{\hat{P}}[\hat{Y}(\mathbb{1}_B\circ\hat{Z})]=E_{\hat{P}}[g_B(\hat{Y},\hat{Z})]=E_P[g_B(Y,Z)]=E_P[Y(\mathbb{1}_B\circ Z)]\\ =\ &E_P[Y\mathbb{1}_A]=E_P[X\mathbb{1}_A]=E_P[X(\mathbb{1}_B\circ Z)]=E_P[g_B(X,Z)]=E_{\hat{P}}[g_B(\hat{X},\hat{Z})]\\ =\ &E_{\hat{P}}[\hat{X}(\mathbb{1}_B\circ\hat{Z})]=E_{\hat{P}}[\hat{X}\mathbb{1}_{\hat{A}}],\end{aligned}$$

which shows the result. □

Tightness of Sequences of Càdlàg Processes

If we endow the Skorokhod space $\mathbb{D}(\mathbb{R}^d)$ of $\mathbb{R}^d$-valued càdlàg functions on $[0,\infty)$ with the Skorokhod metric as explained in Appendix B, then it is a Polish space. So by Prokhorov's theorem any set of probability measures on $\mathbb{D}(\mathbb{R}^d)$ is relatively compact if and only if it is tight. So the classical concept to show weak convergence applies, i.e. for a sequence of distributions on $\mathbb{D}(\mathbb{R}^d)$ (or equivalently for the convergence in distribution of a sequence of stochastic processes with càdlàg paths) one first shows tightness of the sequence of distributions and then one shows that all cluster points coincide.

We first prove a general result concerning "joint tightness" of sequences of random variables. For a metric space (S,d), endowed with the Borel-σ-algebra $\mathcal{B}(S,d)$, we denote by $\mathbb{P}(S)$ the space of probability measures on $(S,\mathcal{B})$.

Lemma 1.42 *Let (S^1,d^1) and (S^2,d^2) be metric spaces. If $(\mu^n)_{n\in\mathbb{N}}$ and $(\nu^n)_{n\in\mathbb{N}}$ are tight sequences in $\mathbb{P}(S^1)$ and $\mathbb{P}(S^2)$, respectively, then the sequence $(\mu^n\otimes\nu^n)_{n\in\mathbb{N}}$ in $\mathbb{P}(S^1\times S^2)$ is tight.*

Proof. Let $\varepsilon > 0$. Then there exist compact sets $K^1_\varepsilon, K^2_\varepsilon$ such that $\mu^n(K^1_\varepsilon) \geq 1 - \frac{\varepsilon}{2}$ and $\nu^n(K^2_\varepsilon) \geq 1 - \frac{\varepsilon}{2}$ for all $n \in \mathbb{N}$, so that for $K_\varepsilon := K^1_\varepsilon \times K^2_\varepsilon$ we have $(\mu^n \otimes \nu^n)(K_\varepsilon) \geq 1 - \varepsilon$ for all $n \in \mathbb{N}$. □

The next result is concerned with tightness of sequences of càdlàg processes. Recall that in the Skorokhod space we have the following modulus of continuity. For $N \in \mathbb{N}$, $\vartheta > 0$, and $\alpha \in \mathbb{D}(\mathbb{R}^d)$ we define

$$w'_N(\alpha, \theta) = \inf \left\{ \max_{i \leq r} w(\alpha; [t_{i-1}, t_i)) \colon r \in \mathbb{N}, 0 = t_0 < \cdots < t_r = N, \inf_{i<r}(t_i - t_{i-1}) \geq \vartheta \right\},$$

where $w(\alpha; I) = \sup_{s,t \in I} |\alpha(s) - \alpha(t)|$ for an interval $I \subseteq \mathbb{R}_+$.

Theorem 1.43 *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of càdlàg processes. Then $(\mathcal{L}(X^n|P^n))_{n\in\mathbb{N}}$ is tight if and only if the following two conditions hold:*

(i) for all $N \in \mathbb{N}$, $\varepsilon > 0$ there exist $n_0 \in \mathbb{N}$, $K > 0$ such that for all $n \geq n_0$

$$P^n \left[\sup_{t \leq N} |X^n_t| > K \right] \leq \varepsilon,$$

(ii) for all $N \in \mathbb{N}$, $\varepsilon > 0$, $\delta > 0$ there exist $n_0 \in \mathbb{N}$, $\vartheta > 0$ such that for all $n \geq n_0$

$$P^n \left[w'_N(X^n, \vartheta) \geq \delta \right] \leq \varepsilon.$$

Proof. cf. Jacod and Shiryaev (1987), Theorem VI.3.21. □

Definition 1.44 Let X and Y be increasing processes. We say that Y *strongly dominates* X, denoted by $X \prec Y$, if $Y - X$ is an increasing process. ◇

In the setting of discrete-time processes which are piecewise constant between deterministic jump times t_k, it is clear that $X \prec Y$ if and only if $\Delta X_{t_k} \leq \Delta Y_{t_k}$ for all k. For the next theorem recall that we denote by $\mathrm{Var}(A^n) = \int |dA^n|$ the total variation process of A^n.

Theorem 1.45 *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of locally square-integrable $\mathbb{F}^n$-semimartingales with canonical decompositions $X^n = X^n_0 + M^n + A^n$. Suppose that $(\mathcal{L}(X^n_0|P^n))_{n\in\mathbb{N}}$ is tight, and that for every $n \in \mathbb{N}$ there exists an increasing $\mathbb{F}^n$-predictable process G^n such that*

$$\langle M^n \rangle + \mathrm{Var}(A^n) \prec G^n,$$

and $\mathcal{L}(G^n|P^n) \xrightarrow{w} \mathcal{L}(G|P)$ for some P-a.s. continuous process G. Then $(\mathcal{L}(X^n|P^n))_{n\in\mathbb{N}}$ is tight.

Proof. cf. Jacod, Mémin and Métivier (1983), Theorem 7.1 with condition C1. □

Convergence Results for Sequences of Semimartingales

If the characteristics of a sequence of semimartingales are known one can show convergence in distribution via convergence of the characteristics. Since for general semimartingales this approach requires additional assumptions like the choice of the probability space and the uniqueness of a corresponding martingale problem, we state these results in the special case where the semimartingales are supposed to have independent increments (but are not necessarily Lévy processes). In view of the results of the next subsection this will be sufficient for our purposes.

To that end we define the following class of functions:

$$\mathcal{C}(\mathbb{R}^d) := \left\{ f \in \mathcal{C}_b(\mathbb{R}^d) \;\middle|\; \exists \varepsilon > 0 \ \forall x \in U_\varepsilon(0) \ f(x) = 0, \text{ and } \lim_{|x|\to\infty} f(x) \text{ exists} \right\},$$

where $\mathcal{C}_b(\mathbb{R}^d)$ is the class of all continuous and bounded functions $f\colon \mathbb{R}^d \to \mathbb{R}$.

Theorem 1.46 *Let X^n, X be $\mathbb{R}^d$-valued processes with independent increments and characteristics B^n, C^n, ν^n and B, C, ν, respectively. Let $\tilde{C}^n$ and $\tilde{C}$ be the modified second characteristics of X^n and X, respectively, assume that X has no fixed times of discontinuity, and let D be a dense subset of $\mathbb{R}_+$. Then $X^n \xrightarrow{\mathcal{L}} X$ if and only if the following three conditions hold:*

(i) $\sup_{s\le t} |B^n_s - B_s| \to 0$ *for all* $t \ge 0$,

(ii) $|\tilde{C}^n_t - \tilde{C}_t| \to 0$ *for all* $t \in D$,

(iii) $g * \nu^n_t \to g * \nu_t$ *for all* $t \in D, g \in \mathcal{C}(\mathbb{R}^d)$.

Proof. cf. Jacod and Shiryaev (1987), Theorem VII.3.4. □

Lemma 1.47 *Let $(X^n)_{n\in\mathbb{N}}$ and $(Y^n)_{n\in\mathbb{N}}$ be two sequences of $\mathbb{R}^d$-valued càdlàg processes. Suppose $X^n \xrightarrow{\mathcal{L}} X$ and $(Y^n)_{n\in\mathbb{N}}$ satisfies*

$$\forall N > 0, \quad \forall \varepsilon > 0 \quad \lim_{n\to\infty} P^n\left[\sup_{t\le N} |Y^n_t| > \varepsilon\right] = 0.$$

Then $X^n + Y^n \xrightarrow{\mathcal{L}} X$.

Proof. cf. Jacod and Shiryaev (1987), Lemma VI.3.31. □

Convergence of Stochastic Integrals and Stochastic Differential Equations

In many cases the semimartingales under consideration are stochastic integrals or solutions of stochastic differential equations driven by a converging sequence of semimartingales. Then one is faced with the question whether the convergence carries over to these new processes. The discussion over this issue dates back to Wong and Zakai (1965) and has received growing interest for obvious reasons. Słominski (1989), Jakubowski, Mémin and Pagès (1989) and Kurtz and Protter (1991) established sufficient conditions for the convergence of stochastic integrals and solutions of stochastic differential equations in terms of uniform tightness of the converging processes, which have the drawback that they are not easy to formulate and sometimes hard to verify. Duffie and Protter (1992) introduce the notion of *goodness* of a sequence of semimartingales and state simple (but not very general) sufficient conditions. See Kurtz and Protter (1996) for more general results.

For a sequence of stochastic bases $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, P^n)$ let $(X^n)_{n\in\mathbb{N}}$ and $(H^n)_{n\in\mathbb{N}}$ be sequences of càdlàg processes where each X^n is an $\mathbb{R}^d$-valued $(\mathbb{F}^n, P^n)$-semimartingale and H^n is $\mathbb{F}^n$-adapted and takes values in $\mathbb{R}^{d'\times d}$. Furthermore on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$ let X be a càdlàg d-dimensional semimartingale and H an adapted càdlàg process with values in $\mathbb{R}^{d'\times d}$. Recall that the total variation process of a process A of finite variation is denoted by $\mathrm{Var}(A) = \int |dA|$.

Definition 1.48 A sequence $(X^n)_{n\in\mathbb{N}}$ of semimartingales is called *good (with respect to $(P^n)_{n\in\mathbb{N}}$ and P)* if for any sequence $(H^n)_{n\in\mathbb{N}}$ the convergence of $\mathcal{L}(X^n, H^n|P^n) \xrightarrow{w} \mathcal{L}(X, H|P)$ implies convergence of $\mathcal{L}(X^n, H^n, \int H^n_- \, dX^n|P^n) \xrightarrow{w} \mathcal{L}(X, H, \int H_- \, dX|P)$. ◇

Proposition 1.49 *Let $(X^n)_{n\in\mathbb{N}}$ be good and suppose $\mathcal{L}(X^n, H^n|P^n) \xrightarrow{w} \mathcal{L}(X, H|P)$. Then $(\int H^n_- \, dX^n)_{n\in\mathbb{N}}$ is also good.*

Proof. cf. Duffie and Protter (1992), Proposition 4.1. □

Theorem 1.50 *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of semimartingales with decompositions $X^n = M^n + A^n$. If $(X^n)_{n\in\mathbb{N}}$ has uniformly bounded jumps and if both $(E^n\big[[M^n, M^n]_T\big])_{n\in\mathbb{N}}$ and $(E^n\left[\mathrm{Var}(A^n)_T\right])_{n\in\mathbb{N}}$ are bounded, then $(X^n)_{n\in\mathbb{N}}$ is good.*

Proof. cf. Duffie and Protter (1992), Theorem 4.1. □

The following (slight) generalization of Duffie and Protter (1992), Example 6.2 shows that multidimensional independent binomial tree models are good.

Example 1.51 Let for all $n \in \mathbb{N}$ $(\xi^n_k)_{k\in\mathbb{N}}$ be a sequence of d-dimensional uniformly bounded i.i.d. random variables and define the sequence $(X^n)_{n\in\mathbb{N}}$ of d-dimensional semimartingales by

$$X^n_t = \sum_{k=1}^{\left[\frac{nt}{T}\right]} \xi^n_k, \quad t \in [0, T],$$

and suppose that $\mathbb{F}^n$ is generated by X^n. If $E_{P^n}[\xi_1^n] = \mathcal{O}\left(\frac{1}{n}\right)$ and $\mathrm{Cov}_{P^n}\left(\xi_1^{ni}, \xi_1^{nj}\right) = \mathcal{O}\left(\frac{1}{n}\right)$, $i, j \in \{1, \ldots, d\}$, then $(X^n)_{n\in\mathbb{N}}$ is good.

Proof. Clearly the decomposition of X^n is given by $X^n = M^n + A^n$ with

$$\begin{aligned} M_t^n &= \sum_{k=1}^{\left[\frac{nt}{T}\right]} \xi_k^n - \left[\frac{nt}{T}\right] E_{P^n}[\xi_1^n], \\ A_t^n &= \left[\frac{nt}{T}\right] E_{P^n}[\xi_1^n]. \end{aligned}$$

Then the quadratic covariation of M^n is given by

$$[M^{ni}, M^{nj}]_t = \sum_{s\le t} \Delta M_s^{ni} \Delta M_s^{nj} = \sum_{k=1}^{\left[\frac{nt}{T}\right]} \left(\xi_k^{ni} - E_{P^n}\left[\xi_1^{ni}\right]\right)\left(\xi_k^{nj} - E_{P^n}\left[\xi_1^{nj}\right]\right),$$

so that $E_{P^n}\left[[M^{ni}, M^{nj}]_T\right] = n\,\mathrm{Cov}_{P^n}\left(\xi_1^{ni}, \xi_1^{nj}\right)$, which is bounded in n by assumption. Furthermore the total variation process $\mathrm{Var}(A^n)$ of A^n is deterministic and given by

$$\mathrm{Var}(A^n)_t = \sum_{s\le t} |\Delta A_s^n| = \left[\frac{nt}{T}\right] |E_{P^n}[\xi_1^n]|,$$

hence $\mathrm{Var}(A^n)_T = n\,|E_{P^n}[\xi_1^n]|$, which is also bounded in n by assumption, so that Theorem 1.50 yields the result, since the jumps of X^n are just the ξ_k^n's, which are uniformly bounded by assumption. □

We next provide a sufficient condition for the convergence of solutions of stochastic differential equations.

Theorem 1.52 *Let $(X^n)_{n\in\mathbb{N}}$ be good, let X be a semimartingale, and let $f\colon \mathbb{R}_+ \times \mathbb{R}^{d'} \to \mathbb{R}^{d'\times d}$ satisfy*

(i) $y \mapsto f(t, y)$ is Lipschitz, uniformly in t,

(ii) $t \mapsto f(t, y)$ is left-continuous with right limits, for all y.

Furthermore let Y^n and Y be the (unique) solutions of

$$\begin{aligned} dY_t^n &= f(t, Y_{t-}^n)\, dX_t^n, \quad Y_0^n \in \mathbb{R}^{d'} \\ dY_t &= f(t, Y_{t-})\, dX_t, \quad Y_0 \in \mathbb{R}^{d'}. \end{aligned}$$

If $X^n \xrightarrow{\mathcal{L}} X$, then $(Y^n, X^n) \xrightarrow{\mathcal{L}} (Y, X)$.

Proof. cf. Duffie and Protter (1992), Theorem 4.4. □

The last theorem has the following drawback. If the driving process is a point process taking values in, say, $\mathbb{Z}^d$, then the solution of a stochastic differential equation usually also takes values in a discrete subset of, say, $\mathbb{R}^{d'}$. So the requirement on f to be Lipschitz is not often met, since f is usually only defined on this discrete subset of $\mathbb{R}^{d'}$. However, the solution process of a stochastic differential equation driven by a point process is constant between jump times of the driving process, a fact which opens new perspectives concerning convergence results on solutions of stochastic differential equation driven by point processes.

Theorem 1.53 *Let $(X^n)_{n\in\mathbb{N}}$ be a sequence of $\mathbb{Z}^d$-valued processes and let X be a point process taking values in $\mathbb{Z}^d$. Let $f\colon \mathbb{Z}^{d'} \to \mathbb{Z}^{d'\times d}$ and denote by Y and Y^n the solutions to the stochastic differential equations*

$$\begin{aligned} dY_t &= f(Y_{t-})\ dX_t, \quad Y_0 = y_0 \in \mathbb{Z}^{d'}, \\ dY_t^n &= f(Y_{t-}^n)\ dX_t^n, \quad Y_0^n = y_0 \in \mathbb{Z}^{d'}. \end{aligned}$$

Consider in addition a sequence $(U^n)_{n\in\mathbb{N}}$ of $\mathbb{R}^{d''}$-valued semimartingales and a continuous $\mathbb{R}^{d''}$-valued semimartingale U, such that $(U^n, X^n) \xrightarrow{\mathcal{L}} (U, X)$. Then

$$(U^n, X^n, Y^n) \xrightarrow{\mathcal{L}} (U, X, Y).$$

Proof. Denote by $\tau_0 = 0$ and $\tau_k = \inf\left\{t > \tau_{k-1} \mid |\Delta X_t| > \frac{1}{2}\right\}$ the jump times of X, and by τ_k^n those of X^n. Then Y is given by

$$Y_t = \sum_{k=0}^{\infty} y_k\ \mathbb{1}_{[\![\tau_k,\tau_{k+1}[\![},$$

where $(y_k)_{k\in\mathbb{N}}$ is a sequence of $\mathbb{Z}^{d'}$ valued random variables, recursively defined by $y_0 \in \mathbb{Z}^{d'}$ and

$$y_k = y_{k-1} + f(y_{k-1})\ \Delta X_{\tau_k}.$$

A further recursion shows that for all $k \in \mathbb{N}$ we have

$$y_k = y_0 + f_k(\Delta X_{\tau_1}, \dots, \Delta X_{\tau_k})$$

for some function $f_k\colon (\mathbb{Z}^d)^k \to \mathbb{Z}^{d'}$. With the same arguments we get for all $n \in \mathbb{N}$

$$Y_t^n = \sum_{k=0}^{\infty} y_k^n\ \mathbb{1}_{[\![\tau_k^n,\tau_{k+1}^n[\![}$$

with

$$y_k^n = y_0 + f_k(\Delta X_{\tau_1}^n, \dots, \Delta X_{\tau_k}^n);$$

note that we have here the same functions f_k as for Y. So we have $(U^n, X^n, Y^n) = \Psi(U^n, X^n)$ for $\Psi(\alpha, \beta) = (\alpha, \Phi(\beta))$, where Φ is the Skorokhod-continuous function from Proposition B.8. Together with Corollary B.3 we thus have that Ψ is continuous on $\mathbb{D}(\mathbb{R}^{d''} \times \mathbb{Z}^d)$ in all points (α, β) such that α is continuous. Therefore $(U^n, X^n, Y^n) \xrightarrow{\mathcal{L}} (U, X, Y)$ follows from the continuous mapping theorem, because U is continuous. □

Contiguity of Sequences of Probability Measures

In the remaining part of this section we are concerned with the following question: When does a change of measure preserve convergence in distribution of a sequence of stochastic processes? We state the most important assertions from Jacod and Shiryaev (1987). To that end let $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n)_{n\in\mathbb{N}}$ be filtered measurable spaces with $\mathcal{F}^n = \mathcal{F}^n_{\infty-} = \bigvee_{t\geq 0} \mathcal{F}^n_t$ and endow each $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n)$ with two probability measures P^n and Q^n.

Definition 1.54 a) The sequence $(Q^n)_{n\in\mathbb{N}}$ is called *contiguous* to the sequence $(P^n)_{n\in\mathbb{N}}$, denoted by $(Q^n)_{n\in\mathbb{N}} \lhd (P^n)_{n\in\mathbb{N}}$, if $P^n(A^n) \stackrel{n\to\infty}{\longrightarrow} 0$ implies $Q^n(A^n) \stackrel{n\to\infty}{\longrightarrow} 0$ for all sequences $(A^n)_{n\in\mathbb{N}}$ with $A^n \in \mathcal{F}^n$.

b) We denote by P^n_t and Q^n_t the restrictions of P^n and Q^n to $\mathcal{F}^n_t$, and we call $(Q^n)_{n\in\mathbb{N}}$ *locally contiguous* to $(P^n)_{n\in\mathbb{N}}$, if $(Q^n_t)_{n\in\mathbb{N}} \lhd (P^n_t)_{n\in\mathbb{N}}$ for all $t \geq 0$. In this case we also write $(Q^n)_{n\in\mathbb{N}} \stackrel{\text{loc}}{\lhd} (P^n)_{n\in\mathbb{N}}$. ◇

Note that the definition of contiguity does not require or imply the existence of limiting measures P and Q. In order to state a condition for contiguity which involves densities, recall that for two measures P, Q and some measure R such that $P \ll R$, $Q \ll R$ with densities z^P, z^Q, respectively, the *Hellinger integral of order* $\alpha \in (0,1)$ is defined as

$$H(\alpha; P, Q) = E_R\left[\left(z^P\right)^\alpha \left(z^Q\right)^{1-\alpha}\right],$$

and it can be shown that $H(\alpha; P, Q)$ is independent of the choice of R. Thus if $Q \ll P$ with density $Z = \frac{dQ}{dP}$, one may choose $R = P$, and the Hellinger integral becomes

$$H(\alpha; P, Q) = E_P\left[Z^{1-\alpha}\right].$$

For details on Hellinger integrals see Jacod and Shiryaev (1987), Chapter IV.

Lemma 1.55 *Denote by P^n_t and Q^n_t the restrictions of P^n and Q^n to $\mathcal{F}^n_t$. Then the following statements are equivalent:*

(i) $(Q^n)_{n\in\mathbb{N}} \stackrel{\text{loc}}{\lhd} (P^n)_{n\in\mathbb{N}}$.

(ii) $\lim_{\alpha\downarrow 0} \liminf_{n\to\infty} H(\alpha; P^n_t, Q^n_t) = 1$ *for all* $t \geq 0$.

Proof. cf. Jacod and Shiryaev (1987), Lemma V.1.6. □

We now come to the main result for our purposes. It answers the question raised above concerning the preservation of convergence in distribution under a change of measure. To that end let $(X^n)_{n\in\mathbb{N}}$ be a sequence of d-dimensional semimartingales, each defined on $(\Omega^n, \mathcal{F}^n, \mathbb{F}^n, P^n)$ and X a d-dimensional semimartingale on $(\Omega, \mathcal{F}, \mathbb{F}, P)$. Furthermore let $Q^n \stackrel{\text{loc}}{\ll} P^n$ for all $n \in \mathbb{N}$ and $Q \stackrel{\text{loc}}{\ll} P$ with density processes Z^n and Z, respectively.

Theorem 1.56 *Assume* $\mathcal{L}((X^n, Z^n)|P^n) \xrightarrow{w} \mathcal{L}((X,Z)|P)$ *and* $(Q^n)_{n\in\mathbb{N}} \overset{\text{loc}}{\triangleleft} (P^n)_{n\in\mathbb{N}}$. *Then* $\mathcal{L}((X^n, Z^n)|Q^n) \xrightarrow{w} \mathcal{L}((X,Z)|Q)$.

Proof. From Jacod and Shiryaev (1987), Theorem X.3.3 it follows that under the above assumptions $\mathcal{L}((X^n, Z^n)|Q^n) \xrightarrow{w} \mu^Q$, where $\mu^Q \ll \mu^P := \mathcal{L}((X,Z)|P)$ is a probability measure on $\big(\mathbb{D}(\mathbb{R}^{d+1}), \mathcal{B}(\mathbb{D}(\mathbb{R}^{d+1}))\big)$ with density process z, such that z is the last component of the coordinate process on $\mathbb{D}(\mathbb{R}^{d+1})$. So it remains to identify μ^Q with $\mathcal{L}((X,Z)|Q)$. However this immediately follows from the transformation of the integrals. For $A \in \mathcal{D}_t(\mathbb{R}^{d+1})$ (for the notation cf. Section 1.1) we have

$$\begin{aligned}
\mu^Q(A) &= \int_{\mathbb{D}(\mathbb{R}^{d+1})} z_t \mathbb{1}_A \, d\mu^P \\
&= \int_\Omega (z_t \mathbb{1}_A) \circ (X,Z) \, dP \\
&= \int_\Omega (Z_t \mathbb{1}_{\{(X,Z)\in A\}}) \, dP \\
&= E_P\left[Z_t \mathbb{1}_{\{(X,Z)\in A\}}\right] \\
&= Q[(X,Z) \in A],
\end{aligned}$$

so $\mu^Q = \mathcal{L}((X,Z)|Q)$ on $\mathcal{D}_t(\mathbb{R}^{d+1})$ for all $t \geq 0$, and since $\mathcal{B}(\mathbb{D}) = \mathcal{D}(\mathbb{R}^{d+1}) = \bigvee_{t\geq 0} \mathcal{D}_t(\mathbb{R}^{d+1})$, we have $\mu^Q = \mathcal{L}((X,Z)|Q)$ on $\mathcal{D}(\mathbb{R}^{d+1})$. $\square$

Remark 1.57 Intuitively it might seem clear that joint convergence of (X,Z) under P^n should imply convergence of X under Q^n, so that it seems redundant to require contiguity of $(Q^n)_{n\in\mathbb{N}}$ with respect to $(P^n)_{n\in\mathbb{N}}$ in Theorem 1.56. However the intuitive argument that

$$E_{Q^n}\left[f(X^n)\right] = E_{P^n}\left[Z^n f(X^n)\right] \overset{n\to\infty}{\longrightarrow} E_P\left[Z f(X)\right] = E_Q\left[f(X)\right]$$

for Z bounded (or at least integrable) works only for $Q^n \sim P^n$ and $Q \sim P$ (then $Z^n = Z^n_\infty$ and $Z = Z_\infty$ above are the usual densities) or in the case of a finite time horizon T (choose $Z^n = Z^n_T$ and $Z = Z_T$ above). One case where the argument does work in an infinite time horizon setting is the case where Z is *continuous*. Then it is possible to apply the continuous mapping theorem and LeCam's first lemma (cf. Jacod and Shiryaev (1987), Corollary V.1.12). In the general case where we have no global relation between Q^n and P^n or Q and P, contiguity is indispensable. $\diamond$

Part II

PRESERVATION OF THE LÉVY PROPERTY UNDER AN OPTIMAL CHANGE OF MEASURE

(This part is an extended and adapted version of Esche and Schweizer (2003))

Chapter 2

Girsanov Quantities and Density Processes

Here we cite the most important results from Jacod and Shiryaev (1987), Chapter III, concerning Girsanov's theorem and the explicit computation of density processes of absolutely continuous probability measures. The chapter mainly serves the purpose of preparation for the main result of Part II, namely the preservation of the Lévy property under the entropy-minimizing martingale measure for a Lévy process, which is presented in Chapter 3.

The basic idea is as follows. Let X be a semimartingale on some stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$. Then it is well known that X remains a semimartingale on $(\Omega, \mathcal{F}, \mathbb{F}, Q)$ where Q is locally absolutely continuous to P. This change of measure can be described by two quantities β and Y, called *Girsanov quantities*, in the sense that the density process Z of Q with respect to P can be expressed via β and Y. These results are summarized in Sections 2.1 and 2.2. In addition we give in Section 2.3 a procedure how one can construct a density process (and thus a locally absolutely continuous measure $\bar{Q}$) from given Girsanov quantities $\bar{\beta}$ and $\bar{Y}$ under certain regularity conditions on $\bar{\beta}$ and $\bar{Y}$. This may be viewed as a converse to Girsanov's theorem.

2.1 Girsanov's Theorem and Girsanov Quantities

In this section we recall Girsanov's theorem for semimartingales, and we introduce the quantities by which one can describe the effect that a change of measure has on the semimartingale characteristics, the so called *Girsanov quantities*. We give examples concerning Lévy processes and Esscher measures.

Let us start with the following useful theorem about stochastic exponentials.

Theorem 2.1 *Let X be a real-valued semimartingale and consider the stochastic differential equation*

$$dZ = Z_- \, dX, \quad Z_0 = 1.$$

This equation has a unique (up to indistinguishability) càdlàg adapted solution, called the stochastic exponential of X, *which is a semimartingale and is denoted by $\mathcal{E}(X)$. Explicitely,*

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}\langle X^c\rangle_t\right) \prod_{s\le t}(1+\Delta X_s)e^{-\Delta X_s}.$$

If we define $\tau := \inf\{t \ge 0 \colon \Delta X_t = -1\}$, then $\mathcal{E}(X) \neq 0$ on $[\![0,\tau[\![$, $\mathcal{E}(X)_- \neq 0$ on $[\![0,\tau]\!]$ and $\mathcal{E}(X) = 0$ on $[\![\tau,\infty[\![$.

Proof. cf. Jacod and Shiryaev (1987), Theorem I.4.61. □

Theorem 2.2 (GIRSANOV'S THEOREM) *Let X be an $\mathbb{R}^d$-valued semimartingale with P-characteristics (B^P, C^P, ν^P) relative to a truncation function h, let c, A be the processes of the "nice version" from Theorem 1.26, and let $Q \overset{\text{loc}}{\ll} P$. Then there exist a predictable function $Y \ge 0$ and a predictable process $\beta = (\beta^i)_{1\le i\le d}$ satisfying*

$$|h\cdot(Y-1)| * \nu^P_t < \infty, \quad \int_0^t |c_s\beta_s| \; dA_s < \infty \text{ and } \int_0^t \beta_s^{\mathrm{tr}} c_s \beta_s \; dA_s < \infty$$

Q-a.s. for all $t \in [0,T]$ and such that the Q-characteristics (B^Q, C^Q, ν^Q) of X are given by

$$\begin{aligned} B_t^{Q,i} &= B_t^{P,i} + \int_0^t (c_s\beta_s)^i \; dA_s + h^i\cdot(Y-1) * \nu_t^P, \quad 1 \le i \le d, \\ C_t^Q &= C_t^P, \\ \nu^Q(dt,dx) &= Y(t,x)\; \nu^P(dt,dx). \end{aligned}$$

Proof. cf. Jacod and Shiryaev (1987), Theorem III.3.24. □

Definition 2.3 The quantities β and Y from Theorem 2.2 are called the *Girsanov quantities of Q with respect to P relative to* X or simply Girsanov quantities of Q if there is no ambiguity about P and the semimartingale involved. ◇

Remark 2.4 Intuitively, Y describes how the jump distributions of X change when we pass from P to Q, and β together with Y determines the change in drift. C^P describes the P-quadratic variation of the continuous part of X and is therefore invariant under an absolutely continuous change of measure. Note, however, that the Girsanov quantities are not unique: From the uniqueness of ν^P and ν^Q we only get uniqueness of $Y(\omega; s, x)$ on supp $\nu^P(\omega)$, and with this and the uniqueness of B^P and B^Q we only get uniqueness of $c\beta$ for fixed c and A. However we can choose the following *nice versions* of Y and β.

To begin with it is obviously possible to choose Y such that $Y(\omega; s, x) = 1$ identically for $(s, x) \notin \text{supp } \nu^P(\omega)$. Note that since ν^P does not charge $\{0\}$, this implies $Y(\omega; s, 0) = 1$ identically.

Now β_s is unique only in the case where c_s is regular. If c_s is degenerate, it is possible to choose β in the following way, where we consider only the case where c is constant in time and deterministic.

Let $\text{rank}(c) = r < d$ and let λ^j be the eigenvalues of c, numbered such that $\lambda^j = 0$ for $j > r$. Since c is nonnegative definite, there exists a diagonal matrix $\tilde{c}$ with $\tilde{c}^{jj} = \lambda^j$ and an orthogonal matrix S such that $c = S\tilde{c}S^{\text{tr}}$. Now let β be any Girsanov quantity and set ${}^S\beta = S^{\text{tr}}\beta$. We write $c\beta = S\tilde{c}S^{\text{tr}}\beta = S\tilde{c}\ {}^S\beta$ and since $\tilde{c}$ is diagonal with $\tilde{c}^{jj} = 0$ for $j > r$, we can set ${}^S\beta^j := 0$ for $j > r$ without changing $c\beta$. So if we set $\gamma^j := {}^S\beta^j = (S^{\text{tr}}\beta)^j$ for $j \le r$ and $\gamma^j := 0$ for $j > r$ and then define $\tilde{\beta} = S\gamma$, we get a new predictable process $\tilde{\beta}$ with $c\tilde{\beta} = c\beta$ and $(S^{\text{tr}}\tilde{\beta})^j = 0$ for $j > r$. Moreover, $\tilde{\beta}$ with these properties is unique. In fact, from $c\tilde{\beta} = c\tilde{\beta}'$ it follows that

$$S\tilde{c}S^{\text{tr}}\tilde{\beta} = c\tilde{\beta} = c\tilde{\beta}' = S\tilde{c}S^{\text{tr}}\tilde{\beta}',$$

and thus, since S is regular,

$$\tilde{c}S^{\text{tr}}\tilde{\beta} = \tilde{c}S^{\text{tr}}\tilde{\beta}',$$

which implies

$$(S^{\text{tr}}\tilde{\beta})^j = (S^{\text{tr}}\tilde{\beta}')^j \quad \text{for } j \le r,$$

since $\tilde{c}$ is diagonal and $c^{jj} \neq 0$ for $j \le r$. Finally, since $(S^{\text{tr}}\tilde{\beta})^j = 0 = (S^{\text{tr}}\tilde{\beta}')^j$ for $j > r$ by construction, we get $S^{\text{tr}}\tilde{\beta} = S^{\text{tr}}\tilde{\beta}'$ and thus $\tilde{\beta} = \tilde{\beta}'$.

We will assume throughout Part II that Y and β are chosen as above, since these considerations simplify some arguments. ◇

Example 2.5 The following example highlights the dependence of the Girsanov quantities on the semimartingale involved. For simplicity we consider the finite time interval $[0, T]$.

We start with the following result from Goll and Kallsen (2000), Lemma A.8. Let $\tilde{L}$ be a one-dimensional P-Lévy process with P-Lévy characteristics $(\tilde{b}, \tilde{c}, \tilde{K})$ relative to some truncation function h. Then $e^{\tilde{L}} = \mathcal{E}(L)$ for some Lévy process L with P-Lévy characteristics (b, c, K), where

$$(2.1) \qquad \begin{cases} b &= \tilde{b} + \frac{1}{2}\tilde{c} + \int_{\mathbb{R}} (h(J(x)) - h(x))\ \tilde{K}(dx) \\ c &= \tilde{c} \\ K(dx) &= \tilde{K} \circ J^{-1}(dx) \end{cases}$$

for $J(x) = e^x - 1$. Note that sometimes the process L is called *stochastic logarithm* of $X = e^{\tilde{L}}$, cf. Kallsen and Shiryaev (2002). Conversely if L is a P-Lévy process with P-Lévy characteristics (b, c, K) and such that $\mathcal{E}(L)$ is strictly positive, then $\mathcal{E}(L) = e^{\tilde{L}}$ for some Lévy process $\tilde{L}$ with P-Lévy characteristics $(\tilde{b}, \tilde{c}, \tilde{K})$, where

$$\left\{ \begin{array}{rcl} \tilde{b} & = & b - \frac{1}{2}c + \int_{\mathbb{R}} \big(h(J^{-1}(x)) - h(x)\big) \; K(dx) \\ \tilde{c} & = & c \\ \tilde{K}(dx) & = & K \circ J(dx). \end{array} \right. \tag{2.2}$$

Now let $u \in \mathbb{R}$ be such that $E_P[\exp(u\tilde{L}_T)] < \infty$ and define $Q \sim P$ by

$$\frac{dQ}{dP} = \frac{\exp(u\tilde{L}_T)}{E_P\left[\exp(u\tilde{L}_T)\right]}.$$

Such a measure Q is called *Esscher measure* for $\tilde{L}$ and it can be shown that $\tilde{L}$ is a Lévy process under Q and that the Girsanov quantities of Q with respect to P relative to $\tilde{L}$ are $\beta^{\tilde{L}} = u$ and $Y^{\tilde{L}}(x) = e^{ux}$ (cf. Propositions 4.2 and 4.5). Let us now determine the Girsanov quantities of Q with respect to P relative to L. By Theorem 2.1 we have

$$\begin{array}{rcl} \tilde{L}_t & = & L_t - \dfrac{1}{2}\langle L^c\rangle_t + \displaystyle\sum_{s\le t} (\log(1+\Delta L_s) - \Delta L_s) \\ & = & L_t - \dfrac{1}{2}ct + (\log(1+x) - x) * \mu^L_t, \end{array}$$

or the other way round,

$$L_t = \tilde{L}_t + \frac{1}{2}ct + (e^x - 1 - x) * \mu^{\tilde{L}}_t,$$

where μ^L and $\mu^{\tilde{L}}$ denote the jump measures of L and $\tilde{L}$, respectively. Note that since these transformations of L and $\tilde{L}$ are invariant under absolutely continuous changes of measure, L is a Lévy process if and only if $\tilde{L}$ is a Lévy process. Now by Theorem 2.2 the Q-Lévy characteristics of $\tilde{L}$ are given by

$$\left\{ \begin{array}{rcl} \tilde{b}^Q & = & \tilde{b} + \tilde{c}u + \int_{\mathbb{R}} h(x)\,(e^{ux} - 1) \; \tilde{K}(dx) \\ \tilde{c}^Q & = & c \\ \tilde{K}^Q(dx) & = & e^{ux} \; \tilde{K}(dx), \end{array} \right. \tag{2.3}$$

and thus by (2.1), the Q-Lévy characteristics of L are

$$\left\{ \begin{array}{rcl} b^Q & = & \tilde{b}^Q + \frac{1}{2}\tilde{c}^Q + \int_{\mathbb{R}}(h(J(x)) - h(x)) \; \tilde{K}^Q(dx) \\ c^Q & = & \tilde{c}^Q \\ K^Q(dx) & = & \tilde{K}^Q \circ J^{-1}(dx). \end{array} \right. \tag{2.4}$$

In order to find the Girsanov quantities of Q with respect to P relative to L, we now express the Q-Lévy characteristics of L in terms of the P-Lévy characteristics of L. First we have

$J^{-1}(x) = \log(1+x)$, thus for $A \in \mathcal{B}(\mathbb{R})$

$$
\begin{aligned}
\int_{\mathbb{R}} \mathbb{1}_A(x)\ K^Q(dx) &= \int_{\mathbb{R}} \mathbb{1}_A(x)\ \tilde{K}^Q \circ J^{-1}(dx) \\
&= \int_{\mathbb{R}} \mathbb{1}_A(J(x))\ \tilde{K}^Q(dx) \\
&= \int_{\mathbb{R}} \mathbb{1}_A(J(x)) e^{ux}\ \tilde{K}(dx) \\
&= \int_{\mathbb{R}} \mathbb{1}_A(x) e^{uJ^{-1}(x)}\ \tilde{K} \circ J^{-1}(dx) \\
&= \int_{\mathbb{R}} \mathbb{1}_A(x)(1+x)^u\ K(dx),
\end{aligned}
$$

and therefore $K^Q(dx) = (1+x)^u\ K(dx)$. Concerning the computation of b^Q we get by (2.4), (2.3) and (2.2)

$$
\begin{aligned}
b^Q &= \tilde{b} + cu + \int_{\mathbb{R}} h(x)\left(e^{ux} - 1\right)\ \tilde{K}(dx) + \frac{1}{2}c + \int_{\mathbb{R}} (h(J(x)) - h(x))\ \tilde{K}^Q(dx) \\
&= b - \frac{1}{2}c + \int_{\mathbb{R}} \big(h(J^{-1}(x)) - h(x)\big)\ K(dx) + cu + \\
&\quad + \int_{\mathbb{R}} h(x)\left(e^{ux} - 1\right)\ \tilde{K}(dx) + \frac{1}{2}c + \int_{\mathbb{R}} (h(J(x)) - h(x))\ \tilde{K}^Q(dx) \\
&= b + cu + \int_{\mathbb{R}} h(J^{-1}(x))\left((1+x)^u - 1\right)\ K(dx) + \\
&\quad + \int_{\mathbb{R}} \big(h(x) - h(J^{-1}(x))\big)\left((1+x)^u - 1\right)\ K(dx) \\
&= b + cu + \int_{\mathbb{R}} h(x)\left((1+x)^u - 1\right)\ K(dx) \\
&= b + c\beta^L + \int_{\mathbb{R}} h(x)\left(Y^L(x) - 1\right)\ K(dx),
\end{aligned}
$$

thus the Girsanov quantities of Q relative to L are given by $\beta^L = u$ and $Y^L(x) = (1+x)^u$. So when we pass from P to Q, the change of drift is the same for $\tilde{L}$ and L, however the change of the jump distributions is different. In particular, an Esscher measure for $\tilde{L}$ is in general not an Esscher measure for L and vice versa; this is a result already stated in Chan (1999), compare the remarks on pp. 523f. there. ◇

Example 2.6 Let W be a standard P-Brownian motion and let A be a $d \times d$-matrix. Set $X := A\,W$ and take $Q \ll P$ with Girsanov quantities β and Y relative to X. Then $X^c = X$, so $\langle X \rangle_t = AA^{\mathrm{tr}}t$, and the P-characteristics of X are given by

$$
\begin{aligned}
B_t^P &= 0 \\
C_t^P &= ct \\
\nu^P(dt, dx) &= 0,
\end{aligned}
$$

where $c := AA^{\mathrm{tr}}$ is the covariance matrix of X. By Theorem 2.2 we get the Q-characteristics of X by

$$\begin{aligned} B_t^{Q,i} &= \int_0^t (c\beta_s)^i \, ds \\ C_t^Q &= ct \\ \nu^Q(dt, dx) &= 0. \end{aligned}$$

Note that X remains a Lévy process under Q if and only if β is deterministic and independent of time. In this case, X is (a linear transformation of standard) Brownian motion with constant drift, i.e. $X_t = A\, W_t^Q + c\beta t$.

In the general case we see that X is still Brownian motion with drift B^Q under Q, which means that in this case, the distribution of X (under Q) is still determined by the Q-characteristics of X, even though X is not a Q-Lévy process. More precisely, we have that $V := X - B^Q \stackrel{d}{=} A\, W^Q$, where W^Q is standard Q-Brownian motion. Indeed, X is a special semimartingale under Q, so from the definition of the characteristics, B^Q is the predictable finite variation part in the canonical decomposition of X under Q. This implies that V is a local Q-martingale, and since X and V are continuous, V is the continuous martingale part of X under Q. Then, again from the definition of the characteristics, $\langle V \rangle_t = ct$ under Q, so like in Theorem 1.36 we have $V \stackrel{d}{=} A\, W^Q$. ◇

2.2 Explicit Computation of the Density Process in the Lévy Case

Let us now consider the case where X has the weak property of predictable representation (in the sense of He, Wang and Yan (1992), Definition 13.13; in Jacod and Shiryaev (1987), Definition III.4.22 and Corollary III.4.27, this is called "all local martingales have the representation property, relative to X"). As we will see, in this case we can even specify the density process of Q. In order to do this, we need some preparation, where we follow Jacod and Shiryaev (1987), Section III.5a, rather closely.

Let X be a semimartingale with P-characteristics (B^P, C^P, ν^P) and μ^X the jump measure of X. Let $Q \stackrel{\mathrm{loc}}{\ll} P$ with Girsanov quantities β and Y. We define the following stochastic processes:

$$\begin{aligned} a_t(\omega) &= \nu^P(\omega, \{t\} \times \mathbb{R}^d), \\ \hat{Y}_t(\omega) &= \begin{cases} \int Y(\omega, t, x)\, \nu^P(\omega, \{t\} \times dx) & \text{if this integral converges} \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

If ν^P is deterministic, the set $\{t\colon a_t > 0\}$ is the set of all fixed times of discontinuity of X (cf. Jacod and Shiryaev (1987), Theorem II.4.15), so in this case $\hat{Y}$ denotes the intensities under Q of the jumps of X at fixed times of discontinuity.

For the predictable time σ, defined by

$$\sigma = \inf\{t \geq 0 \mid \text{either } \hat{Y}_t > 1, \text{ or } a_t = 1 \text{ and } \hat{Y}_t < 1\},$$

we introduce a generalized (i.e. not necessarily finite-valued) increasing process H by

$$\begin{aligned} H_t &= \int_0^t (\beta_s^{\text{tr}} c \beta_s) \mathbb{1}_{[\![0,\sigma]\!]} \, dA_s + \left(1 - \sqrt{Y(\omega, s, x)}\right)^2 \mathbb{1}_{[\![0,\sigma]\!]} * \nu_t^P + \\ &\quad + \sum_{s \leq t} \left(\sqrt{1 - a_s} - \sqrt{1 - \hat{Y}_s}\right)^2 \mathbb{1}_{\{s<\sigma\}} \end{aligned}$$

and set

$$\begin{aligned} \tau_n &= \inf\{t \geq 0 \mid H_t \geq n\}, \\ \Delta &= [\![0, \sigma[\![\cap \left(\bigcup_n [\![0, \tau_n]\!]\right). \end{aligned}$$

Proposition 2.7 *There is a process N, unique up to P-indistinguishability on Δ, such that for every stopping time S with $[\![0, S]\!] \subseteq \Delta$, the stopped process N^S is the local P-martingale*

$$N^S = \int \beta \mathbb{1}_{[\![0,S]\!]} \, dX^c + \left(Y - 1 + \frac{\hat{Y} - a}{1 - a} \mathbb{1}_{\{a<1\}}\right) \mathbb{1}_{[\![0,S]\!]} * (\mu^X - \nu^P).$$

Proof. cf. Jacod and Shiryaev (1987), Proposition III.5.10. □

Theorem 2.8 *Let $Q \overset{\text{loc}}{\ll} P$ with density process Z and assume that X has the weak property of predictable representation. Then*

$$Z_t = \begin{cases} Z_0 \; \mathcal{E}(N)_t & \text{if } (\omega, t) \in \Delta \\ 0 & \text{otherwise,} \end{cases}$$

where N and Δ are defined as above.

Proof. cf. Jacod and Shiryaev (1987), Theorem III.5.19. □

From here we derive the following useful result concerning the explicit computation of density processes in a filtration $\mathbb{F}$ generated by a P-Lévy process L. Let μ^L be the jump measure of L, ν^P its (deterministic) P-compensator and let (b, c, K) be the P-Lévy characteristics of L.

Corollary 2.9 *Let $\mathbb{F}$ be the filtration generated by a P-Lévy process L and let $Q \overset{\text{loc}}{\sim} P$ with Girsanov quantities β, Y. We define the stochastic process N^Q by*

$$N_t^Q = \int_0^t \beta_s \, dL_s^c + (Y - 1) * (\mu^L - \nu^P)_t.$$

Then the density process of Q with respect to P is given by

$$Z_t^Q = \mathcal{E}(N^Q)_t, \quad t \in [0, \infty).$$

Proof. Note that L has the weak property of predictable representation (cf. Jacod and Shiryaev (1987), Section II.6 and Theorem III.4.34.). So Theorem 2.8 implies $Z = Z_0\, \mathcal{E}(N)\, \mathbb{1}_\Delta$, with N as in Proposition 2.7. Furthermore, $Q \overset{\text{loc}}{\sim} P$ implies $Z_t > 0$ P-a.s. for all t, so $\Delta = [\![0, \infty[\![$ (cf. Jacod and Shiryaev (1987), Corollary III.5.22). Then for any stopping time S we have $[\![0, S]\!] \subseteq \Delta$, so N is already a local P-martingale (take $S = +\infty$), and $Z = \mathcal{E}(N)$. (Note that $\mathcal{F}_0$ is trivial, so $Z_0 = 1$.) In addition, a Lévy process has no fixed times of discontinuity, so $\nu(\omega, \{t\} \times \mathbb{R}) = 0$ for all t, hence a and $\hat{Y}$ vanish in the Lévy case. This yields the claimed form of N^Q. □

2.3 Construction of Equivalent Measures in Terms of Given Girsanov Quantities

In Theorem 2.2 we have seen a "parametrization" of the set of absolutely continuous measures $Q \overset{\text{loc}}{\ll} P$ in terms of β and Y. However "parametrization" is a slight abuse of terminology because although we get Girsanov quantities β and Y for a *given measure* Q it is a different issue to determine whether for *given quantities* β and Y there exists a measure $Q \overset{\text{loc}}{\ll} P$ with these Girsanov quantities. If we define N^Q from β and Y as in Corollary 2.9, then the natural candidate for Q should have $Z^Q := \mathcal{E}(N^Q)$ as density process and then we can run through the above arguments in reverse order. But to make sure that the natural candidate for Q is actually a probability measure, we need to know at least that Z^Q is a true P-martingale. Theorem 2.11 will be useful to deal with this first issue.

There is a second issue here which concerns the existence of $Q \overset{\text{loc}}{\ll} P$. If we are given a nonnegative P-martingale Z and a fixed *finite* time horizon $[0, T]$, we can always define $Q \ll P$ on $\mathcal{F}_T$ by $dQ = Z_T\, dP$, and then of course $Q \overset{\text{loc}}{\ll} P$ when we restrict our attention to the interval $[0, T]$. For an arbitrary stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and *infinite* time horizon there is no such construction for Q on all of $\mathcal{F}$ in general (unless Z is uniformly integrable or defined on $[0, \infty]$, of course). However there is the following positive result, if we work on the path space. Recall that we choose all filtrations to be right-continuous.

Theorem 2.10 *Let $(\Omega, \mathcal{F}) = \big(\mathbb{D}([0, \infty), \mathbb{R}^d), \mathcal{B}(\mathbb{D})\big)$, let P be a probability measure on $(\Omega, \mathcal{F})$ and let $\mathbb{F}^{\mathbb{D}} = (\mathcal{F}_t^{\mathbb{D}})_{t \geq 0}$ be the P-augmentation of the filtration generated by the coordinate process on Ω. Let Z be a nonnegative $(P, \mathbb{F}^{\mathbb{D}})$-martingale with càdlàg paths and $Z_0 = 1$. Then there exists a probability measure $Q \overset{\text{loc}}{\ll} P$ on $(\Omega, \mathcal{F})$ with $dQ = Z_t\, dP$ on $\mathcal{F}_t^{\mathbb{D}}$ for all $t \geq 0$.*

Proof. cf. Kallenberg (1997), Lemma 16.18. □

The discussion around Theorem 2.10 makes it clear that for the explicit construction of locally absolutely continuous measures by means of a given density process in the case of an infinite time horizon, it will be useful and necessary to confine ourselves to the Skorokhod space as our stochastic basis.

We now state the following criteria for the stochastic exponential of a local martingale M to be a true martingale. Note that in the case of a continuous local martingale M, this criterion reduces to the well-known Novikov criterion.

Theorem 2.11 *Let M be a real-valued local P-martingale on $[0,\infty]$ with $M_0 = 0$ and $\Delta M > -1$ P-a.s. If the process A, defined by*

$$A_t := \frac{1}{2}\langle M^c\rangle_t + \sum_{s\le t} \left((1+\Delta M_s)\log(1+\Delta M_s) - \Delta M_s\right), \tag{2.5}$$

admits a predictable P-compensator B with $E_P[\exp B_\infty] < \infty$, then $\mathcal{E}(M)$ is a uniformly integrable P-martingale and $\mathcal{E}(M)_\infty > 0$ P-a.s.

Proof. cf. Lepingle and Mémin (1978), Théorème III.1. □

Corollary 2.12 *Let M and A be as in Theorem 2.11 and suppose A admits a P-compensator B with $E_P[\exp B_t] < \infty$ for all $t \geq 0$. Then $\mathcal{E}(M)$ is a P-martingale on $[0,\infty)$ and $\mathcal{E}(M)_t > 0$ P-a.s. for all $t \geq 0$.*

Proof. The claim follows if we show that $E_P[\mathcal{E}(M)_\tau] = E_P[\mathcal{E}(M)_0] = 1$ for every bounded stopping time τ. So let $\tau \leq T$ for some deterministic $T > 0$. Then

$$\mathcal{E}(M)_\tau = \mathcal{E}(M)^T_\tau = \mathcal{E}(M^T)_\tau.$$

Now M^T is again a local P-martingale with $M^T_0 = 0$ and $\Delta M^T > -1$, and the corresponding process $A(M^T) = A^T$ admits a P-compensator $B(M^T) = B^T$ which satisfies $E_P\left[\exp B(M^T)_\infty\right] = E_P[\exp B_T] < \infty$ by assumption. So Theorem 2.11 yields that $\mathcal{E}(M^T)$ is a uniformly integrable P-martingale, and thus $E_P[\mathcal{E}(M)_\tau] = E_P\left[\mathcal{E}(M^T)_\tau\right] = 1$.

The positivity of $\mathcal{E}(M)_t$ for all t follows with the same argument, because for $t \leq T$ we have $\mathcal{E}(M)_t = \mathcal{E}(M^T)_t = E_P\left[\mathcal{E}(M^T)_\infty \middle| \mathcal{F}_t\right] > 0$ P-a.s. because $\mathcal{E}(M^T)_\infty = \mathcal{E}(M)_T > 0$ P-a.s.
□

Let us now apply this criterion in order to construct from given quantities β and Y strictly positive martingales which will serve as density processes of locally equivalent measures under certain conditions. We start with the following technical lemma.

Lemma 2.13 *Define the functions $f, g\colon [0,\infty) \to \mathbb{R}$ by*

$$\begin{aligned} f(y) &= \begin{cases} y\log y - (y-1) & \text{if } y > 0 \\ 1 & \text{if } y = 0, \end{cases} \\ g(y) &= (1-\sqrt{y})^2. \end{aligned}$$

Then f and g are convex, and $0 \leq g(y) \leq f(y)$ for all $y \geq 0$.

Proof. We have $f''(y) = \frac{1}{y} > 0$ and $g''(y) = \frac{1}{2y\sqrt{y}} > 0$ for all $y > 0$, thus f and g are convex. That $g \geq 0$ is immediate. We show $k := f - g \geq 0$. In fact, $k \in C^\infty$ on $(0, \infty)$ with $k'(y) = \log y - 1 + \frac{1}{\sqrt{y}}$ and $k''(y) = \frac{1}{y} - \frac{1}{2y\sqrt{y}}$. Now $y = \frac{1}{4}$ is the only zero of k''; hence k' has at most one local extremum and thus at most two zeros. This in turn means that k has at most two local extrema. But $y = 1$ is a local minimum with $k(1) = 0$, and since $k(0) = 0$ the second extremum of k is a local maximum somewhere between 0 and 1. Hence $k(y) \geq 0$ for all $y \geq 0$. □

Proposition 2.14 *Let L be a P-Lévy process with jump measure μ^L and Lévy characteristics (b, c, K). Let $\bar{\beta}$ be a predictable process and $\bar{Y} > 0$ be a predictable function and suppose*

$$E_P \left[\exp \left(\int_0^t \left(\frac{1}{2} \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s + \int_{\mathbb{R}^d} f(\bar{Y}(s,x)) \; K(dx) \right) \; ds \right) \right] < \infty \tag{2.6}$$

for all $t \geq 0$. Then $\bar{Y} - 1$ is integrable with respect to $\mu^L - \nu^P$ and $\bar{Z} := \mathcal{E}(\bar{N})$ with

$$\bar{N}_t := \int_0^t \bar{\beta}_s \; dL_s^c + (\bar{Y} - 1) * (\mu^L - \nu^P)_t, \quad t \geq 0, \tag{2.7}$$

is a strictly positive P-martingale on $[0, \infty)$.

Proof. Note that by Lemma 2.13 and the fact that $f(\bar{Y}) * \nu^P \geq 0$ P-a.s. we have for all $t \geq 0$

$$g(\bar{Y}) * \nu_t^P \leq f(\bar{Y}) * \nu_t^P \leq \exp \left(\int_0^t \left(\frac{1}{2} \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s + \int_{\mathbb{R}^d} f(\bar{Y}(s,x)) \; K(dx) \right) \; ds \right) \tag{2.8}$$

because $z \leq e^z$ for all z. So $(1 - \sqrt{\bar{Y}})^2 * \nu^P$ is locally P-integrable by (2.6), and from Theorem 1.22 c) we get the integrability of $\bar{Y} - 1$ with respect to $\mu^L - \nu^P$. Furthermore $\int \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s \; ds$ is also locally P-integrable by (2.6), so that $\bar{\beta}$ is integrable with respect to L^c. With this, $\bar{N}$ is well-defined, and the second claim will follow from Corollary 2.12, if $\bar{N}$ is a local P-martingale with $\Delta \bar{N} > -1$ P-a.s. and if the process in (2.5) with $M := \bar{N}$ admits a compensator B with $E[\exp B_t] < \infty$ for all $t \geq 0$.

By definition of the continuous martingale part and the stochastic integral with respect to a compensated measure, L^c and $(\bar{Y} - 1) * (\mu^L - \nu^P)$ are local P-martingales, so altogether $\bar{N}$ is a local P-martingale. Moreover, (2.7) is the decomposition of the local martingale $\bar{N}$ into a continuous and a purely discontinuous local martingale. Hence part b) of Definition 1.20 shows that

$$\Delta \bar{N}_t = \Delta \left((\bar{Y} - 1) * (\mu^L - \nu^P)_s \right) = (\bar{Y}(s, \Delta L_s) - 1) \mathbb{1}_{\{\Delta L_t \neq 0\}} > -1$$

P-a.s. since $\bar{Y} > 0$.

To find the P-compensator B of A for $\bar{N}$, note that $\bar{N}^c = \int \bar{\beta} \; dL^c$, so by Theorem 1.36 $\langle \bar{N}^c \rangle_t = \langle \int \bar{\beta} \; dL^c \rangle_t = \int_0^t \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s \; ds$. Therefore in Corollary 2.12 we have for $M = \bar{N}$

$$\begin{aligned} A_t &= \frac{1}{2} \int_0^t \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s \; ds + \sum_{s \leq t} \left(\bar{Y}(s, \Delta L_s) \; \log \bar{Y}(s, \Delta L_s) - \bar{Y}(s, \Delta L_s) + 1 \right) \mathbb{1}_{\{\Delta L_s \neq 0\}} \\ &= \frac{1}{2} \int_0^t \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s \; ds + f(\bar{Y}) * \mu_t^L. \end{aligned}$$

Now $|f(\bar{Y})| * \nu_t^P = f(\bar{Y}) * \nu_t^P$ is P-integrable for all $t \geq 0$ by (2.8), so Theorem 1.21 implies integrability of $f(\bar{Y})$ with respect to $\mu^L - \nu^P$ and that $f(\bar{Y}) * (\mu^L - \nu^P) = f(\bar{Y}) * \mu^L - f(\bar{Y}) * \nu^P$. This yields that

$$B_t = \frac{1}{2} \int_0^t \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s \, ds + f(\bar{Y}) * \nu_t^P$$

is the P-compensator of A, and we have $E[\exp B_t] < \infty$ for all $t \geq 0$ by assumption. □

If $\bar{\beta}$ and $\bar{Y}$ are deterministic and independent of time, the conditions of Proposition 2.14 are particularly easy to verify:

Corollary 2.15 *Let L be a P-Lévy process with jump measure μ^L and Lévy characteristics (b, c, K) and let $\bar{\beta} \in \mathbb{R}^d$ be a constant and $\bar{Y} \colon \mathbb{R}^d \to (0, \infty)$ a measurable function. If*

$$\int_{\mathbb{R}^d} f(\bar{Y}(x)) \; K(dx) < \infty,$$

then $\bar{Z} := \mathcal{E}(\bar{N})$ with

$$\bar{N}_t := \bar{\beta}^{\mathrm{tr}} L_t^c + (\bar{Y} - 1) * (\mu^L - \nu^P)_t, \quad t \geq 0,$$

is a strictly positive P-martingale on $[0, \infty)$.

Proof. For $\bar{\beta}$ and $\bar{Y}$ deterministic and independent of time, (2.6) reduces to

$$\begin{aligned} E_P &\left[\exp \left(\int_0^t \left(\frac{1}{2} \bar{\beta}_s^{\mathrm{tr}} c \bar{\beta}_s + \int_{\mathbb{R}^d} f(\bar{Y}(s, x)) \; K(dx) \right) \; ds \right) \right] = \\ &= \quad \exp \left(\left(\bar{\beta}^{\mathrm{tr}} c \bar{\beta} + \int_{\mathbb{R}^d} f(\bar{Y}(x)) \; K(dx) \right) t \right), \end{aligned}$$

which is finite by assumption. □

We now turn to the following question: Does a measure $\bar{Q}$ with a density process constructed from given quantities $\bar{\beta}$ and $\bar{Y}$ have these quantities as Girsanov quantities? If such a $\bar{Q}$ exists (as is the case if $\Omega = \mathbb{D}$), we have the following result.

Proposition 2.16 *Let L be a P-Lévy process with jump measure μ^L, P-compensator ν^P and Lévy characteristics (b, c, K) relative to a fixed truncation function h. Let $\bar{\beta}, \bar{Y}$ be as in Proposition 2.14 and such that $\bar{Y} - 1$ is integrable with respect to $\mu^L - \nu^P$, and define*

$$\bar{N} = \int \bar{\beta} \; dL^c + (\bar{Y} - 1) * (\mu^L - \nu^P).$$

Suppose there is a probability measure $\bar{Q} \overset{\mathrm{loc}}{\sim} P$ with density process $Z^{\bar{Q}} = \mathcal{E}(\bar{N})$. Then $\bar{\beta}$ and $\bar{Y}$ are the Girsanov quantities of $\bar{Q}$.

Proof. The density process of the given measure $\bar{Q}$ is a strictly positive P-martingale, so $\bar{Z} := \mathcal{E}(\bar{N})$ is also a strictly positive P-martingale. On the other hand, Theorem 2.2 gives us a predictable function $\hat{Y} \geq 0$ and a predictable process $\hat{\beta} = (\hat{\beta}^i)_{1 \leq i \leq d}$ with $\int_0^t \hat{\beta}_s^{\mathrm{tr}} c \hat{\beta}_s \, ds < \infty$ and $|h(x)(\hat{Y}(s,x) - 1)| * \nu_t^P < \infty$, and by Corollary 2.9 we know that $Z^{\bar{Q}} = \mathcal{E}(N^{\bar{Q}})$ with

$$N_t^{\bar{Q}} = \int_0^t \hat{\beta}_s \, dL_s^c + (\hat{Y} - 1) * (\mu^L - \nu^P)_t.$$

So since $\mathcal{E}(N^{\bar{Q}}) = Z^{\bar{Q}} = \bar{Z} = \mathcal{E}(\bar{N}) > 0$, we have $N^{\bar{Q}} = \bar{N}$, or equivalently,

$$\int_0^t (\bar{\beta}_s - \hat{\beta}_s) \, dL_s^c = (\hat{Y} - \bar{Y}) * (\mu^L - \nu^P)_t, \quad t \geq 0.$$

But the left-hand side above is a continuous local P-martingale and the right-hand side is a purely discontinuous local P-martingale. So both sides vanish and we show that this implies $\hat{\beta} = \bar{\beta}$ and $\hat{Y} = \bar{Y}$.

Since $\int (\bar{\beta} - \hat{\beta}) \, dL^c \equiv 0$ is obviously a continuous local P-martingale with vanishing quadratic variation, we have

$$\int_0^t (\bar{\beta}_s - \hat{\beta}_s)^{\mathrm{tr}} c (\bar{\beta}_s - \hat{\beta}_s) \, ds = 0 \quad P\text{-a.s. for all } t \geq 0.$$

Now the integrand is P-a.s. nonnegative so that

$$(\bar{\beta}_s - \hat{\beta}_s)^{\mathrm{tr}} c (\bar{\beta}_s - \hat{\beta}_s) = 0 \tag{2.9}$$

P-a.s. for all $s \geq 0$. Recall that $\bar{\beta}$ and $\hat{\beta}$ are chosen as in Remark 2.4, i.e. we have $(S^{\mathrm{tr}} \beta)^j = 0$ for $\beta \in \{\hat{\beta}, \bar{\beta}\}$ and for $j > \mathrm{rank}(c)$. (S is the unique orthogonal matrix such that $c = S \tilde{c} S^{\mathrm{tr}}$, where $\tilde{c}$ is the diagonal matrix of the eigenvalues of c with $\tilde{c}^{jj} = 0$ for $j > \mathrm{rank}(c)$.) With this in mind (2.9) is equivalent to

$$(\bar{\beta}_s - \hat{\beta}_s)^{\mathrm{tr}} S \tilde{c} S^{\mathrm{tr}} (\bar{\beta}_s - \hat{\beta}_s) = 0,$$

which implies $(S^{\mathrm{tr}}(\bar{\beta}_s - \hat{\beta}_s))^j = 0$ for $j \leq \mathrm{rank}(c)$, and since $(S^{\mathrm{tr}} \hat{\beta}_s)^j = 0 = (S^{\mathrm{tr}} \bar{\beta}_s)^j$ for $j > \mathrm{rank}(c)$ by construction, we have $S^{\mathrm{tr}}(\bar{\beta}_s - \hat{\beta}_s) = 0$, and thus $\bar{\beta}_s = \hat{\beta}_s$ P-a.s. for all $s \geq 0$.

Concerning the equality of $\bar{Y}$ and $\hat{Y}$, we note that $M := (\hat{Y} - \bar{Y}) * (\mu^L - \nu^P) \equiv 0$ is a square-integrable local P-martingale with vanishing variance process $\langle M \rangle$. From Theorem 1.22 a) we then get for all $t \geq 0$

$$0 = \langle M \rangle_t = (\hat{Y} - \bar{Y})^2 * \nu_t^P = \int_0^t \int_{\mathbb{R}^d} (\hat{Y}(s,x) - \bar{Y}(s,x))^2 \, K(dx) ds$$

P-a.s., so that P-a.s. $\hat{Y}(s,x) = \bar{Y}(s,x)$ ν^P-a.e. Thus $\bar{\beta}$ and $\bar{Y}$ are the Girsanov quantities of $\bar{Q}$. □

With the help of Corollary 2.15 and Proposition 2.16 we can now construct (at least on the Skorokhod space if we consider an infinite time horizon) locally equivalent measures which preserve the Lévy structure of L: just choose $\bar{\beta}$ and $\bar{Y}$ deterministic and independent of time. This is a converse of Girsanov's theorem for Lévy processes, and we state the result as a corollary for future reference. An alternative direct construction of $\bar{Q}$ is given in the appendix, see Theorem A.9.

Corollary 2.17 *Let $(\Omega, \mathcal{F}) = \big(\mathbb{D}([0,\infty), \mathbb{R}^d), \mathcal{B}(\mathbb{D})\big)$ and L be the coordinate process on $\mathbb{D}$. Let P be a probability measure on $(\Omega, \mathcal{F})$ and let $\mathbb{F} = \mathbb{F}^L(P)$. Suppose that L is a P-Lévy process and let (b, c, K) be the Lévy characteristics of L relative to a fixed truncation function h. Let $\bar{\beta} \in \mathbb{R}^d$ be a constant and $\bar{Y} : \mathbb{R}^d \to (0, \infty)$ a measurable function with*

$$\int_{\mathbb{R}^d} f(\bar{Y}(x)) \; K(dx) < \infty.$$

Then there exists $\bar{Q} \overset{\text{loc}}{\sim} P$ with Girsanov quantities $\bar{\beta}, \bar{Y}$, and L is a $\bar{Q}$-Lévy process with $\bar{Q}$-Lévy characteristics

$$\begin{aligned} b^{\bar{Q},i} &= b^i + (c\bar{\beta})^i + \int_{\mathbb{R}^d} h^i(x)(\bar{Y}(x) - 1) \; K(dx), \quad 1 \le i \le d, \\ c^{\bar{Q}} &= c, \\ K^{\bar{Q}}(dx) &= \bar{Y}(x) \; K(dx). \end{aligned}$$

In the case of a finite time horizon $[0, T]$ with $\mathcal{F} = \mathcal{F}_T$, the statement remains true for any stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, P)$ and P-Lévy process L, if $\mathbb{F} = \mathbb{F}^L(P)$.

Proof. Corollary 2.15 yields that $\bar{Z} = \mathcal{E}(\bar{N})$ with $\bar{N}_t := \bar{\beta}^{\text{tr}} L^c_t + (\bar{Y} - 1) * (\mu^L - \nu^P)_t$, $t \ge 0$ is a strictly positive P-martingale, so by Theorem 2.10 there exists $\bar{Q} \overset{\text{loc}}{\sim} P$ with density process $\bar{Z}$. (In the case of a finite time horizon $[0, T]$ define $\bar{Q}$ by $d\bar{Q} = \bar{Z}_T \; dP$.) Then by Proposition 2.16 $\bar{\beta}$ and $\bar{Y}$ are the Girsanov quantities of $\bar{Q}$, and we get the special form of the $\bar{Q}$-Lévy characteristics immediately from Theorem 2.2, since $\bar{Y}$ and $\bar{\beta}$ are deterministic and independent of time and the P-characteristics of L are deterministic and linear in time. $\square$

Chapter 3

Preservation of the Lévy Property

We now come to the main result of Part II. Throughout this chapter, let L be a d-dimensional P-Lévy process with respect to the P-augmentation of the filtration generated by L, jump measure μ^L and P-Lévy characteristics (b, c, K) relative to a fixed truncation function h. Define $\nu^P(ds, dx) = dsK(dx)$ and let U be a fixed $d \times d$-matrix. From Corollary 1.35 we know that UL is again a Lévy process, and the aim in this chapter is to show that L remains a Lévy process under the entropy-minimizing martingale measure $Q^E(UL)$ for UL. For simplicity, we write in the sequel $\mathcal{Q}_x^U := \mathcal{Q}_x^U(L)$ for $x \in \{a, e, f, \ell\}$ and $Q^E := Q^E(UL)$. The matrix U will be important for certain applications involving models with stochastic volatility, see Section 4.4.

The idea is the following: We show that relative entropy of some martingale measure $Q \overset{\text{loc}}{\ll} P$ is a convex functional of its Girsanov quantities Y and β, so by Jensen's inequality we obtain a measure Q^ℓ with smaller relative entropy if we pass to deterministic time-independent Girsanov quantities by averaging over ω and t. Furthermore the martingale property of L is characterized by a linear constraint between Y and β, and is therefore preserved by this averaging. Thus we construct for any (local) martingale measure $Q \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U$ a Lévy martingale measure $Q^\ell \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$ with smaller relative entropy than Q. Hence Q^E, if it exists, must be a Lévy martingale measure. Section 3.1 is devoted to the construction of a measure, which preserves the Lévy structure of L and which has smaller relative entropy than a given measure. In Section 3.2 we show how to parametrize (local) martingale measures in terms of Girsanov quantities. Finally in Section 3.3 we show that Q^E is in fact a Lévy martingale measure.

3.1 Improvement of Relative Entropy

In this section we show how one can construct to a given measure $Q \overset{\text{loc}}{\sim} P$ with finite-valued entropy process for each $t \geq 0$ a measure Q^ℓ with the following properties:

(i) L is a Q^ℓ-Lévy process.

(ii) $I_t(Q^\ell|P) \leq I_t(Q|P)$.

Of course the first choice for Q^ℓ should be P itself since L is a P-Lévy process and since $I(P|P) = 0 \leq I(Q|P)$ for all measures Q. But since we want to find the optimal measure in $\mathcal{Q}_a^U$ we construct Q^ℓ in such a way that $Q^\ell \in \mathcal{Q}_a^U$ if $Q \in \mathcal{Q}_a^U$, which is shown in Section 3.3.

In Chapter 2 we have seen that under an absolutely continuous change of measure a P-Lévy process with P-characteristics (B^P, C^P, ν^P), which are deterministic and linear in time, becomes a Q-semimartingale with Q-characteristics (B^Q, C^Q, ν^Q) which are in general not deterministic and linear in time any more. The change of the characteristics when we pass from P to Q can be described by the Girsanov quantities β and Y. Since the characteristics describe drift, volatility and jumps, the idea is to construct a measure Q^ℓ by taking deterministic and time-independent Girsanov quantities β^ℓ and Y^ℓ which leave B^Q and ν^Q unchanged in the mean. This is done by averaging the "cumulated" quantities $\int \beta_s \, ds$ and $\int Y(s,x) \, ds$. Then the Q^ℓ-characteristics of L should be the "average" Q-characteristics.

For the following preparatory lemmas we fix some $Q \overset{\text{loc}}{\sim} P$ with Girsanov quantities β and Y and $I_t(Q|P) < \infty$ for all $t \geq 0$. Let $Z = \mathcal{E}(N)$ be the density process of Q with respect to P, where

$$N_t = \int_0^t \beta_s \, dL_s^c + (Y-1) * (\mu^L - \nu^P)_t$$

by Corollary 2.9. From Lemma 1.11 we know that $Z \log Z$ is a P-submartingale. Recall from Lemma 2.13 that $f, g \colon \mathbb{R}_+ \to \mathbb{R}_+$ with $f(y) = y \log y - (y-1)$ and $g(y) = (1-\sqrt{y})^2$ are convex and that we have $0 \leq g \leq f$.

Lemma 3.1 *The canonical decomposition of the P-submartingale $Z \log Z$ is given by*

$$Z \log Z = M + A \tag{3.1}$$

with

$$M = \int Z_-(1 + \log Z_-) \, dN + (Z_- f(Y)) * (\mu^L - \nu^P)$$

and

$$A = \frac{1}{2} \int Z_- d\langle N^c \rangle + (Z_- f(Y)) * \nu^P.$$

Proof. By integration by parts, we have

$$d(Z \log Z) = Z_- \, d(\log Z) + (\log Z_-) \, dZ + d[Z, \log Z]. \tag{3.2}$$

Now Itô's formula for the explicit form of $\mathcal{E}(N)$ yields

$$\log Z_t = N_t - \frac{1}{2}\langle N^c \rangle_t + \sum_{s \leq t} \left(\log(1 + \Delta N_s) - \Delta N_s \right),$$

where the sum is absolutely convergent for all $t \geq 0$. In fact $|\Delta N_s| > \frac{1}{2}$ only for finitely many $s \leq t$, and for $|x|^2 \leq \frac{1}{2}$ we have $|\log(1+x) - x| \leq \text{const. } |x|^2$. Thus

$$\sum_{s \leq t} |\log(1+\Delta N_s) - \Delta N_s| \mathbb{1}_{\{|\Delta N_s| \leq \frac{1}{2}\}} \leq \text{const.} \sum_{s \leq t} |\Delta N_s|^2 \leq \text{const. } [N]_t < \infty,$$

since $[N]$ is finite-valued (cf. Jacod and Shiryaev (1987), Theorem I.4.47). Furthermore, the process $D := \sum_s (\log(1+\Delta N_s) - \Delta N_s)$ is of finite variation and

$$\log Z = N - \frac{1}{2}\langle N^c \rangle + D.$$

This gives us the $d(\log Z)$-term in (3.2).

Let us now examine the $d[Z, \log Z]$-term in (3.2). We have $dZ = Z_- \, dN$, and thus

$$d[Z, \log Z] = Z_- \, d[N, \log Z] = Z_- \left(d[N] - \frac{1}{2} \, d[N, \langle N^c \rangle] + d[N, D] \right). \tag{3.3}$$

Since $\langle N^c \rangle$ is continuous, $[N, \langle N^c \rangle]$ vanishes, and since D is of finite variation, we have

$$[N, D]_t = \sum_{s \leq t} \Delta N_s \Delta D_s = \sum_{s \leq t} \Delta N_s \left(\log(1+\Delta N_s) - \Delta N_s \right). \tag{3.4}$$

This sum is absolutely convergent, since

$$\sum_{s \leq t} |\Delta N_s \Delta D_s| = \int_0^t |d[N, D]_s| \leq ([N]_t)^{\frac{1}{2}} ([D]_t)^{\frac{1}{2}}$$

by the Kunita-Watanabe inequality. And since $\sum_{s \leq t} (\Delta N_s)^2 \leq [N]_t$ converges as well, we can decompose the sum in (3.4) and get

$$[N, D] = \sum_s \Delta N_s \log(1+\Delta N_s) - \sum_s (\Delta N_s)^2,$$

which at last yields

$$\begin{aligned} [N, \log Z] &= [N] - \sum_s (\Delta N_s)^2 + \sum_s \Delta N_s \log(1+\Delta N_s) \\ &= \langle N^c \rangle + \sum_s \Delta N_s \log(1+\Delta N_s), \end{aligned}$$

or in terms of (3.3)

$$[Z, \log Z] = \int Z_- \, d\langle N^c \rangle + \sum_s Z_{s-} \Delta N_s \log(1+\Delta N_s).$$

Putting all this together and using $dZ = Z_- \, dN$, we finally get

$$\begin{aligned} Z \log Z &= \int Z_- \, dN - \frac{1}{2} \int Z_- \, d\langle N^c \rangle + \int Z_- \, dD + \int Z_- \log Z_- \, dN + \\ &\quad + \int Z_- \, d\langle N^c \rangle + \sum_s Z_{s-} \, \Delta N_s \, \log(1 + \Delta N_s) \\ &= \int Z_-(1 + \log Z_-) \, dN + \frac{1}{2} \int Z_- \, d\langle N^c \rangle + \sum_s Z_{s-} \left(\Delta D_s + \Delta N_s \, \log(1 + \Delta N_s) \right) \\ &= \int Z_-(1 + \log Z_-) \, dN + \frac{1}{2} \int Z_- \, d\langle N^c \rangle + \sum_s Z_{s-} \, f(1 + \Delta N_s). \qquad (3.5) \end{aligned}$$

This is a decomposition of the P-submartingale $Z \log Z = M' + A' + V$, where

$$\begin{aligned} M' &:= \int Z_-(1 + \log Z_-) \, dN \quad \text{is a local } P\text{-martingale}, \\ A' &:= \frac{1}{2} \int Z_- \, d\langle N^c \rangle \quad \text{is continuous and increasing}, \\ V &:= \sum_s Z_{s-} f(1 + \Delta N_s) \quad \text{is increasing}, \end{aligned}$$

since $\Delta N_s > -1$ and $f(y) \geq 0$ for $y \geq 0$. However, V is not predictable so that (3.5) is not the canonical decomposition. So let us further decompose V into a local P-martingale and a predictable process of finite variation. We can write

$$V = Z \log Z - M' - A',$$

where all terms on the right-hand side are locally P-integrable: $Z \log Z$ since $I_t(Q|P) < \infty$ for all $t \geq 0$, M' because it is a local P-martingale, and A' because it is continuous. On the other hand, $\Delta N_s = (Y(s, \Delta L_s) - 1) \mathbb{1}_{\{\Delta L_s \neq 0\}}$ so that $Y(s, 0) = 1$ yields

$$f(1 + \Delta N_s) = f(Y(s, \Delta L_s)) \mathbb{1}_{\{\Delta L_s \neq 0\}}.$$

This in turn gives

$$V = (Z_- f(Y)) * \mu^L = |Z_- f(Y)| * \mu^L,$$

since $Z_- f(Y) \geq 0$. Now U is locally P-integrable, so Theorem 1.21 implies that

$$(Z_- f(Y)) * \mu^L = (Z_- f(Y)) * (\mu^L - \nu^P) + (Z_- f(Y)) * \nu^P.$$

So altogether we have

$$Z \log Z = \left(M' + (Z_- f(Y)) * (\mu^L - \nu^P) \right) + \left(A' + (Z_- f(Y)) * \nu^P \right),$$

and this in fact is the canonical decomposition since the first term is a local P-martingale and the second term is predictable and of finite variation. □

Lemma 3.2 *Let $A = \frac{1}{2}\int Z_- d\langle N^c\rangle + (Z_- f(Y)) * \nu^P =: A' + A''$ be the predictable process of finite variation from the canonical decomposition (3.1) of $Z \log Z$. Then A'_t and A''_t are P-integrable for all $t \geq 0$.*

Proof. We show that A_t is P-integrable for some arbitrary but fixed $t \geq 0$. Then the claim follows from the nonnegativity of A' and A''.

Now $Z \log Z$ is a P-submartingale, and for any fixed t, the family

$$\{Z_\tau \log Z_\tau \mid \tau \leq t \text{ is an } \mathbb{F}\text{-stopping time}\}$$

is uniformly integrable since

$$-e^{-1} \leq Z_\tau \log Z_\tau \leq E_P\left[Z_t \log Z_t \mid \mathcal{F}_\tau\right]$$

and $Z_t \log Z_t \in \mathcal{L}^1(P)$ since $I_t(Q|P) < \infty$. Thus $(Z \log Z)^t$ is a submartingale of class (D), and therefore the increasing process of its Doob-Meyer decomposition is integrable with respect to P. But the uniqueness of the decomposition yields $(Z \log Z)^t = M^t + A^t$, hence $E_P[A_t] = E_P[A^t_\infty] < \infty$. □

Lemma 3.3 *a) The following random variables are Q-integrable for each $t \geq 0$:*

(i) $\int_0^t \beta_s^{\mathrm{tr}} c \beta_s \, ds$

(ii) $\int_0^t |\beta_s^i| \, ds$ *for* $1 \leq i \leq d$

(iii) $f(Y) * \nu_t^P$

(iv) $\int_0^t Y(s,x) \, ds$ *for* $x \in \operatorname{supp} K$.

b) The entropy process of Q with respect to P is given by

$$I_t(Q|P) = E_Q\left[\frac{1}{2}\int_0^t (\beta_s)^{\mathrm{tr}} c \beta_s \, ds + f(Y) * \nu_t^P\right].$$

Proof. a) Fix some arbitrary $t \geq 0$.

(i) From Lemma 3.2 we know that $A'_t = \frac{1}{2}\int_0^t Z_{s-} \, d\langle N^c\rangle_s \in \mathcal{L}^1(P)$. Now, Dellacherie and Meyer (1980), Théorème VI.61 yields

$$E_P\left[\int_0^t Z_{s-} \, d\langle N^c\rangle_s\right] = E_P\left[Z_t \langle N^c\rangle_t\right] = E_Q\left[\langle N^c\rangle_t\right]$$

so that we get

$$\int_0^t \beta_s^{\mathrm{tr}} c \beta_s \, ds = \langle N^c\rangle_t \in \mathcal{L}^1(Q).$$

(Note that $\langle N^c\rangle$ is the quadratic variation of the continuous P-martingale part of N.)

(ii) Let $r = \operatorname{rank}(c)$ and λ^j be the eigenvalues of c, numbered such that $\lambda^j = 0$ exactly for $r < j \leq d$. Choose β as in Remark 2.4, i.e. $\beta_s = S\gamma_s$ with $\gamma_s^j = 0$ for $r < j \leq d$ and S orthogonal such that $c = S\tilde{c}S^{\mathrm{tr}}$ for a diagonal matrix $\tilde{c}$ with $\tilde{c}^{jj} = \lambda^j$. Then

$$\beta_s^{\mathrm{tr}} c \beta_s = (S^{\mathrm{tr}}\beta_s)^{\mathrm{tr}} \tilde{c} (S^{\mathrm{tr}}\beta) = \gamma_s^{\mathrm{tr}} \tilde{c} \gamma_s = \sum_{j=1}^{r} \lambda^j |\gamma_s^j|^2$$

and $\beta_s^i = \sum_{j=1}^r S^{ij}\gamma_s^j$ so that

$$\int_0^t |\beta_s^i| \; ds = \int_0^t \left| \sum_{j=1}^r S^{ij}\gamma_s^j \right| \; ds \leq \sum_{j=1}^r \left|S^{ij}\right| \int_0^t |\gamma_s^j| \; ds.$$

Hence it suffices to show that $\int_0^t |\gamma_s^j| \; ds$ is Q-integrable. In fact,

$$\begin{aligned}
\left(E_Q \left[\frac{1}{t} \int_0^t |\gamma_s^i| \; ds \right] \right)^2 &\leq E_Q \left[\left(\frac{1}{t} \int_0^t |\gamma_s^i| \; ds \right)^2 \right] \\
&\leq E_Q \left[\frac{1}{t} \int_0^t |\gamma_s^i|^2 \; ds \right] \\
&\leq \text{const. } E_Q \left[\int_0^t \sum_{j=1}^r \lambda^j |\gamma_s^j|^2 \; ds \right] \\
&= \text{const. } E_Q \left[\int_0^t \beta_s^{\mathrm{tr}} c \beta_s \; ds \right]
\end{aligned}$$

which is finite by (i). Furthermore $\left| \int_0^T \beta_s \; ds \right| \leq \int_0^T |\beta_s| \; ds$, so $\int_0^T \beta_s \; ds$ is also Q-integrable.

(iii) Again from Lemma 3.2, we have $A_t'' = (Z_- f(Y)) * \nu_t^P \in \mathcal{L}^1(P)$, and as above, using Dellacherie and Meyer (1980), Théorème VI.61, we get

$$\begin{aligned}
E_P[(Z_- f(Y)) * \nu_t^P] &= E_P \left[\int_0^t Z_{s-} \int_{\mathbb{R}^d} f(Y(s,x)) \; K(dx)ds \right] \\
&= E_P \left[Z_t \int_0^t \int_{\mathbb{R}^d} f(Y(s,x)) \; K(dx)ds \right] \\
&= E_Q \left[f(Y) * \nu_t^P \right],
\end{aligned}$$

and thus $f(Y) * \nu_t^P \in \mathcal{L}^1(Q)$.

(iv) From (iii) we know that $E_Q\left[f(Y) * \nu_t^P\right] = E_Q\left[\int_{\mathbb{R}^d \times [0,t]} f(Y(s,x)) \; dsK(dx)\right] < \infty$, so with Fubini's theorem we get for $x \in \operatorname{supp} K$

$$E_Q \left[\int_0^t f(Y(s,x)) \; ds \right] < \infty.$$

Since f is convex, we get by first using Jensen's inequality for Q and then for the uniform distribution on $[0,t]$ that

$$\begin{aligned} f\left(E_Q\left[\frac{1}{t}\int_0^t Y(s,x)\;ds\right]\right) &\le E_Q\left[f\left(\frac{1}{t}\int_0^t Y(s,x)\;ds\right)\right] \\ &\le E_Q\left[\frac{1}{t}\int_0^t f(Y(s,x))\;ds\right] < \infty. \end{aligned}$$

Finally $f(y) < \infty$ implies $y < \infty$, hence the claim.

b) Concerning the explicit form of the entropy process, we note that by Lemma 3.1 the canonical decomposition of the P-submartingale $Z \log Z$ is given by

$$Z \log Z = M + A = M + \frac{1}{2}\int Z_-\;d\langle N^c\rangle + (Z_- f(Y)) * \nu^P.$$

Now for fixed $t \ge 0$ we have already seen that the stopped submartingale $(Z\log Z)^t$ is of class (D), so that M^t is a uniformly integrable P-martingale, and thus

$$I_t(Q|P) = E_P[Z_t \log Z_t] = E_P[A_t] = E_Q\left[\frac{1}{2}\int_0^t (\beta_s)^{\mathrm{tr}} c\beta_s\;ds + f(Y) * \nu_t^P\right]$$

from a) (i) and (iii) and the proofs thereof. □

We are now in a position to state the main result of this section: To any locally equivalent measure Q for L with finite-valued entropy process we can construct a "better" measure Q^ℓ which leaves the characteristics of L unchanged in the mean and under which L is a Lévy process. Here "better" is meant in the sense that Q^ℓ has smaller relative entropy than Q as seen in the next proposition. In order to distinguish between Q and Q^ℓ we denote in the sequel the Girsanov quantities of Q with respect to P by β^Q and Y^Q.

Theorem 3.4 *Let $T > 0$ be fixed and $Q \overset{\mathrm{loc}}{\sim} P$ with $I_T(Q|P) < \infty$. Define*

$$\begin{aligned} Y^\ell(x) &= \frac{1}{T}\,E_Q\left[\int_0^T Y^Q(s,x)\;ds\right], \quad x \in \operatorname{supp} K, \\ \beta^{\ell,i} &= \frac{1}{T}\,E_Q\left[\int_0^T \beta_s^{Q,i}\;ds\right], \quad 1 \le i \le d. \end{aligned}$$

a) There exists a probability measure $Q^\ell \sim P$ on $\mathcal{F}_T$ with Girsanov quantities β^ℓ and Y^ℓ, which satisfies $I_T(Q^\ell|P) \le I_T(Q|P)$, and such that the restriction of L to the interval $[0,T]$ is a Q^ℓ-Lévy process.

b) Let $(\Omega,\mathcal{F}) = \big(\mathbb{D}([0,\infty),\mathbb{R}^d),\mathcal{B}(\mathbb{D})\big)$ and L be the coordinate process on $\mathbb{D}$ and set $\mathbb{F} = \mathbb{F}^L(P)$. Then there exists $Q^\ell \overset{\mathrm{loc}}{\sim} P$ with Girsanov quantities β^ℓ and Y^ℓ, which satisfies $I_T(Q^\ell|P) \le I_T(Q|P)$, and such that L is a Q^ℓ-Lévy process on the interval $[0,\infty)$.

c) Let Q^ℓ be constructed as above. Then $I_T(Q^\ell|P) = I_T(Q|P)$ if and only if $\beta^Q = \beta^\ell$ and $Y^Q(.,x) = Y^\ell(x)$ $P \otimes \lambda_{[0,T]}$-a.s. for all $x \in \operatorname{supp} K$, where $\lambda_{[0,T]}$ denotes the Lebesgue measure on $([0,T], \mathcal{B}([0,T]))$ (i.e. if and only if L is a Q-Lévy process on $[0,T]$).

Proof. Note that β^ℓ and Y^ℓ are well-defined thanks to (the proof of) Lemma 3.3 a). If we show that $\int f(Y^\ell(x))\ K(dx) < \infty$, Corollary 2.17 yields the existence of a measure $Q^\ell \sim P$ on $\mathcal{F}_T$ (or $Q^\ell \overset{\text{loc}}{\sim} P$ in the case where $\Omega = \mathbb{D}$) with Girsanov quantities β^ℓ and Y^ℓ under which L is a Lévy process on $[0,T]$ (or on $[0,\infty)$ in the case where $\Omega = \mathbb{D}$).

Since f is convex, using twice Jensen's inequality and then Fubini's theorem gives

$$\begin{aligned}
\int_{\mathbb{R}^d} f(Y^\ell(x))\ K(dx) &= \int_{\mathbb{R}^d} f\left(E_Q\left[\frac{1}{T}\int_0^T Y^Q(s,x)\ ds\right]\right) K(dx) \\
&\le \int_{\mathbb{R}^d} E_Q\left[\frac{1}{T}\int_0^T f(Y^Q(s,x))\ ds\right] K(dx) \\
&= \frac{1}{T} E_Q\left[f(Y^Q) * \nu_T^P\right] < \infty
\end{aligned}$$

by Lemma 3.3 a).

Concerning the improvement of relative entropy we know from Lemma 3.3 b) that $I(R|P)$ is given by

$$I(R|P) = E_R\left[\frac{1}{2}\int_0^T (\beta_s^R)^{\mathrm{tr}} c \beta_s^R\ ds + f(Y^R) * \nu_T^P\right],$$

for $R \in \{Q, Q^\ell\}$. We show that

(i) $E_Q\left[\int_0^T (\beta_s^Q)^{\mathrm{tr}} c \beta_s^Q\ ds\right] \ge E_{Q^\ell}\left[\int_0^T (\beta^\ell)^{\mathrm{tr}} c \beta^\ell\ ds\right] = T(\beta^\ell)^{\mathrm{tr}} c \beta^\ell$,

(ii) $E_Q\left[f(Y^Q) * \nu_T^P\right] \ge E_{Q^\ell}\left[f(Y^\ell) * \nu_T^P\right] = T\int_{\mathbb{R}^d} f(Y^\ell(x))\ K(dx)$,

with equality if and only if $\beta^Q = \beta^\ell$ and $Y^Q(.,x) = Y^\ell(.,x)$ $P \otimes \lambda_{[0,T]}$-a.s. for all $x \in \operatorname{supp} K$.

We start with (ii). Since f is convex, we have by Jensen's inequality and Fubini's theorem

$$\begin{aligned}
E_Q\left[f(Y^Q) * \nu_T^P\right] &= T\ E_Q\left[\int_{\mathbb{R}^d} \frac{1}{T}\int_0^T f(Y^Q(s,x))\ dsK(dx)\right] \\
&\ge T\ E_Q\left[\int_{\mathbb{R}^d} f\left(\frac{1}{T}\int_0^T Y^Q(s,x)\ ds\right)\ K(dx)\right] \\
&\ge T \int_{\mathbb{R}^d} f\left(E_Q\left[\frac{1}{T}\int_0^T Y^Q(s,x)\ ds\right]\right)\ K(dx) \\
&= T \int_{\mathbb{R}^d} f(Y^\ell(x))\ K(dx) \\
&= f(Y^\ell) * \nu_T^P \\
&= E_{Q^\ell}\left[f(Y^\ell) * \nu_T^P\right],
\end{aligned}$$

with equality if and only if $Y^Q(.,x) = Y^\ell(.,x)$ $P\otimes\lambda_{[0,T]}$-a.s. for all $x \in \operatorname{supp} K$ by Lemma C.6.

Let us now show (i). With the notation of Remark 2.4 we have

$$(\beta_s^Q)^{\mathrm{tr}} c \beta_s^Q = (\beta_s^Q)^{\mathrm{tr}} S\tilde{c}S^{\mathrm{tr}}\beta_s^Q = |\tilde{\gamma}_s|^2$$

with $\tilde{\gamma}_s = \sqrt{\tilde{c}}\gamma = \sqrt{\tilde{c}}S^{\mathrm{tr}}\beta_s^Q$ and $\left(\sqrt{\tilde{c}}\right)^{ij} = \sqrt{\tilde{c}^{ij}}$. By Jensen's inequality for the uniform distribution on $[0,T]$ we have $\frac{1}{T}\int_0^T |\tilde{\gamma}_s|^2\, ds \geq \left|\frac{1}{T}\int_0^T \tilde{\gamma}_s\, ds\right|^2$, which implies

$$\begin{aligned} E_Q\left[\frac{1}{T}\int_0^T (\beta_s^Q)^{\mathrm{tr}} c\beta_s^Q\, ds\right] &= E_Q\left[\frac{1}{T}\int_0^T |\tilde{\gamma}_s|^2\, ds\right] \\ &\geq E_Q\left[\left|\frac{1}{T}\int_0^T \tilde{\gamma}_s\, ds\right|^2\right] \\ &\geq \left|\frac{1}{T}E_Q\left[\int_0^T \tilde{\gamma}_s\, ds\right]\right|^2, \end{aligned}$$

with equality if and only if $\tilde{\gamma}$ is constant $P\otimes\lambda_{[0,T]}$-a.s. (which is the case if and only if β^Q is so) by Lemma C.6. Note that the last estimate uses Jensen's inequality with respect to Q for the d coordinates. However we have

$$\frac{1}{T}E_Q\left[\int_0^T \tilde{\gamma}_s\, ds\right] = \frac{1}{T}E_Q\left[\int_0^T \sqrt{\tilde{c}}S^{\mathrm{tr}}\beta_s^Q\, ds\right] = \sqrt{\tilde{c}}S^{\mathrm{tr}}\frac{1}{T}E_Q\left[\int_0^T \beta_s^Q\, ds\right] = \sqrt{\tilde{c}}S^{\mathrm{tr}}\beta^\ell,$$

and thus

$$\left|\frac{1}{T}E_Q\left[\int_0^T \tilde{\gamma}_s\, ds\right]\right|^2 = (\beta^\ell)^{\mathrm{tr}}S\tilde{c}S^{\mathrm{tr}}\beta^\ell = (\beta^\ell)^{\mathrm{tr}}c\beta^\ell.$$

So altogether we have

$$E_Q\left[\int_0^T (\beta_s^Q)^{\mathrm{tr}}c\beta_s^Q\, ds\right] \geq T(\beta^\ell)^{\mathrm{tr}}c\beta^\ell,$$

with equality if and only if $\beta^Q = \beta^\ell$ $P\otimes\lambda_{[0,T]}$-a.s., hence (i) is proved. □

3.2 Parametrization of Classes of Martingale Measures

Intuitively, a measure $Q \overset{\mathrm{loc}}{\ll} P$ can be described via two quantities β and Y that determine the Q-characteristics of L from those under P as seen in Chapter 2. If L should be a local Q-martingale, we know from Corollary 1.33 that $B^Q + (x - h(x)) * \mu^L$ must be a local Q-martingale, where B^Q is the first Q-characteristic of L. This gives us a relation between β and Y. Hence a martingale measure for L should be determined by a single quantity Y, and in the case where L is a Q-Lévy process, this should further reduce to a non-random function.

Let us now make these ideas more precise. Recall that U is a fixed $d\times d$-matrix and U_h is defined by

$$U_h(x) = Uh(x) - h(Ux), \quad x \in \mathbb{R}^d.$$

We see from Lemma C.3 that $\int_{\mathbb{R}^d} |U_h(x)| \ K(dx) < \infty$ (take $Y \equiv 1$ there), and that $E_Q[|U_h|Y * \nu_t^P] < \infty$ whenever $E_Q[f(Y) * \nu_t^P] < \infty$, which is the case if the entropy process of Q with respect to P is finite-valued, cf. Lemma 3.3 b). Also note that we are actually looking for conditions on β and Y for UL to be a local Q-martingale.

Lemma 3.5 *Let $Q \overset{\text{loc}}{\ll} P$ with Girsanov quantities β and Y and $E_Q[f(Y) * \nu_t^P] < \infty$ for all $t \geq 0$. Then*

a) $|Ux - h(Ux)|Y * \nu_t^P < \infty$ Q-a.s. if and only if $|U(xY - h)| * \nu_t^P < \infty$ Q-a.s.

b) $|Ux - h(Ux)|Y * \nu_t^P$ is Q-integrable if and only if $|U(xY - h)| * \nu_t^P$ is Q-integrable.

Proof. By the triangular inequality we have

$$|Ux - h(Ux)|Y \leq |U(xY - h)| + |Uh \cdot (1 - Y)| + |Uh(x) - h(Ux)|Y \tag{3.6}$$

and

$$|U(xY - h)| \leq |Ux - h(Ux)|Y + |Uh \cdot (1 - Y)| + |Uh(x) - h(Ux)|Y. \tag{3.7}$$

Now $|Uh(x) - h(Ux)|Y * \nu_t^P \leq \text{const.}(t + f(Y) * \nu_t^P)$ by Lemma C.3 and also

$$|Uh \cdot (1 - Y)| * \nu_t^P = |Uh \cdot (Y - 1)| * \nu_t^P \leq \text{const.}\, |h \cdot (Y - 1)| * \nu_t^P \leq \text{const.}(t + f(Y) * \nu_t^P)$$

by Lemma C.5. So the last two terms on the right hand side in (3.6) and (3.7) are Q-integrable by assumption and thus Q-a.s. finite. This yields the claim. □

Lemma 3.6 *Let $Q \overset{\text{loc}}{\ll} P$ with Girsanov quantities β and Y and $E_Q[f(Y) * \nu_t^P] < \infty$ for all $t \geq 0$. Suppose one of the following two conditions holds:*

(i) UL is a local Q-martingale,

(ii) $|U(xY - h)| * \nu_t^P < \infty$ Q-a.s. for all $t \geq 0$.

Then $|Ux - h(Ux)| * \mu^L$ is locally Q-integrable.

Proof. (i) Let UL be a local Q-martingale. Then UL is a special semimartingale under Q and by the fact that $|Ux - h(Ux)| \leq \text{const.}(|x|^2 \wedge |x|)$ we get local Q-integrability of $|Ux - h(Ux)| * \nu^Q$ (and thus of $|Ux - h(Ux)| * \mu^L$) by Proposition 1.30.

(ii) Note that by Lemma 3.5 a) finiteness of $|U(xY - h)| * \nu^P$ is equivalent to finiteness of $|Ux - h(Ux)|Y * \nu^P$. Then $|Ux - h(Ux)| * \nu^Q = |Ux - h(Ux)| \ Y * \nu^P$ is Q-a.s. finite-valued and continuous (recall $\nu^P(ds, dx) = dsK(dx)$) and thus locally Q-integrable. However, this is equivalent to $|Ux - h(Ux)| * \mu^L$ being locally Q-integrable; see Theorem 1.21. □

Proposition 3.7 *Let L be a P-Lévy process with P-Lévy characteristics (b, c, K) relative to some truncation function h. Let $Q \overset{\text{loc}}{\ll} P$ with Girsanov quantities β and Y and suppose $E_Q[f(Y) * \nu_t^P] < \infty$ for all $t \geq 0$. Then UL is a local Q-martingale if and only if $|U(x\, Y(s,x) - h(x))| * \nu_t^P < \infty$ and*

$$U\left(b + c\beta_t + \int_{\mathbb{R}^d} (x\, Y(t,x) - h(x))\ K(dx)\right) = 0 \tag{3.8}$$

Q-a.s. for all $t \geq 0$.

Definition 3.8 Condition (3.8) is called the *martingale condition for UL.* ◇

Note that the martingale condition is independent of the choice of the truncation function. In fact, if we replace h in the definition of the characteristics by some other truncation function h', then b is replaced by $b' = b - \int_{\mathbb{R}^d}(h(x) - h'(x))\ K(dx)$ (cf. Jacod and Shiryaev (1987), Proposition II.2.24), so that (3.8) holds with (b', c, K) relative to h'.

Proof of Proposition 3.7. From Corollary 1.33 we know that a necessary and sufficient condition for the semimartingale UL with P-characteristics $(\tilde{B}, \tilde{C}, \tilde{\nu}^P)$ relative to h to be a local Q-martingale is that $\tilde{B}^Q + (x - h(x)) * \mu^{UL}$ be a local Q-martingale. By Theorem 2.2 the Q-characteristics of L are

$$\begin{aligned} B_t^Q &= bt + \int_0^t c\beta_s\ ds + (h \cdot (Y - 1)) * \nu_t^P \\ C_t^Q &= ct \\ \nu^Q(ds, dx) &= Y(s,x)\ \nu^P(ds, dx) \end{aligned}$$

and by Theorem 1.34 the Q-characteristics of UL are

$$\begin{aligned} \tilde{B}_t^Q &= UB_t^Q - U_h * \nu_t^Q \\ \tilde{C}_t^Q &= UC_t^Q U^{\text{tr}} \\ \tilde{\nu}^Q(A_1 \times A_2) &= \nu^Q\left(A_1 \times U^{-1}(A_2 \setminus \{0\}\right), \quad A_1 \in \mathcal{B}(\mathbb{R}_+), A_2 \in \mathcal{B}(\mathbb{R}^d). \end{aligned}$$

With this the first Q-characteristic of UL is given by

$$\begin{aligned} \tilde{B}_t^Q &= Ubt + \int_0^t Uc\beta_s\ ds + (Uh \cdot (Y - 1)) * \nu_t^P - U_h * \nu_t^Q \\ &= Ubt + \int_0^t Uc\beta_s\ ds + (Uh \cdot (Y - 1) - U_h Y) * \nu_t^P. \end{aligned}$$

We define the process $\tilde{M}$ by

$$\tilde{M}_t := \tilde{B}_t^Q + (x - h(x)) * \mu_t^{UL}$$

so that UL is a local Q-martingale if and only if $\tilde{M}$ is a local Q-martingale by Corollary 1.33. Note that

$$(x - h(x)) * \mu_t^{UL} = (Ux - h(Ux))1\!\!1_{\{|Ux| \neq 0\}} * \mu_t^L = (Ux - h(Ux)) * \mu_t^L,$$

and thus if $|Ux - h(Ux)| * \mu_t^L$ is locally Q-integrable, we can write

$$(x - h(x)) * \mu^{UL} = (Ux - h(Ux)) * (\mu^L - \nu^Q) + (Ux - h(Ux)) * \nu^Q.$$

Using $\nu^Q(ds, dx) = Y(s, x)\ \nu^P(ds, dx) = Y(s, x)\ dsK(dx)$ (recall that L is a P-Lévy process) we then get

$$\begin{aligned}
\tilde{M}_t &= Ubt + \int_0^t Uc\beta_s\ ds + (Uh \cdot (Y-1) - U_h Y) * \nu_t^P + (Ux - h(Ux)) * (\mu^L - \nu^Q)_t \\
&\quad + (Ux - h(Ux)) * \nu_t^Q \\
&= \Big(Uh(x)(Y(s,x) - 1) - U_h(x)Y(s,x) + (Ux - h(Ux))Y(s,x)\Big) * \nu_t^P + Ubt \\
&\quad + \int_0^t Uc\beta_s\ ds + (Ux - h(Ux)) * (\mu^L - \nu^Q)_t \\
&= \int_0^t \int_{\mathbb{R}^d} (U(xY(s,x) - h(x)))\ K(dx)ds + Ubt + \int_0^t Uc\beta_s\ ds \\
&\quad + (Ux - h(Ux)) * (\mu^L - \nu^Q)_t \\
&=: \int_0^t \tilde{m}_s\ ds + (Ux - h(Ux)) * (\mu^L - \nu^Q)_t.
\end{aligned}$$

Observe that $\tilde{m}_t = 0$ Q-a.s. for all $t \geq 0$ is just condition (3.8).

Now suppose UL is a local Q-martingale. Then $|Ux - h(Ux)| * \mu^L$ is locally Q-integrable by Lemma 3.6 and thus finite-valued. This implies $|U(xY - h)| * \nu_t^P < \infty$ Q-a.s. by Lemma 3.5 a). Furthermore $\tilde{M} = \int \tilde{m}_s\ ds + (Ux - h(Ux)) * (\mu^L - \nu^Q)$ is a local Q-martingale. Thus $\int \tilde{m}_s\ ds$ is a continuous local Q-martingale of finite variation and thus $\int \tilde{m}_s\ ds \equiv 0$ Q-a.s., which implies $\tilde{m} \equiv 0$ Q-a.s.

To show the other direction, suppose $|U(xY - h)| * \nu_t^P < \infty$ and that (3.8) is satisfied. $|U(xY - h)| * \nu_t^P < \infty$ implies that $|Ux - h(Ux)| * \mu^L$ is locally Q-integrable by Lemma 3.6, and thus $\tilde{M} = \int \tilde{m}_s\ ds + (Ux - h(Ux)) * (\mu^L - \nu^Q)$. Then (3.8) yields $\tilde{M} = (Ux - h(Ux)) * (\mu^L - \nu^Q)$ which is a local Q-martingale by Theorem 1.21. Hence UL is a local Q-martingale. □

Remark 3.9 Note that if U and c are regular, then (3.8) is equivalent to

$$\beta_t = -c^{-1}\left(b + \int_{\mathbb{R}^d} (x\ Y(t,x) - h(x))\ K(dx)\right) \quad Q\text{-a.s. for all } t \geq 0,$$

where the integral is to be read componentwise. ◇

Proposition 3.7 gives us a "parametrization" of the set $\mathcal{Q}_a^U$ of absolutely continuous local martingale measures for UL in terms of Y. In the case where U is regular, β is then determined via the martingale condition (if c is not regular, choose the nice version of β as in Remark 2.4). Again "parametrization" is meant in the sense that for a *given* martingale measure Q for L we get Girsanov quantities β and Y. However note the result of Corollary 2.17, where we have constructed $\bar{Q} \overset{\text{loc}}{\sim} P$ from given quantities $\bar{\beta}$ and $\bar{Y}$ under the sole condition that $\bar{\beta}$ and $\bar{Y}$ be deterministic and independent of time and that $\bar{Y}$ satisfies $\int f(\bar{Y}(x))\, K(dx) < \infty$. In the case of deterministic quantities, this is equivalent to finite relative entropy. Obviously if $\bar{\beta}$ and $\bar{Y}$ are chosen such that in addition the martingale condition for UL holds, such a measure $\bar{Q}$ is a Lévy martingale measure for UL. So such quantities parametrize exactly the set $\mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$ of all Lévy martingale measures for UL with finite relative entropy.

Example 3.10 Here we show how one can construct Lévy martingale measures in a canonical way, if U and c are regular. So let L be a P-Lévy process with P-Lévy-characteristics (b, c, K). We define the deterministic and time-independent function

$$Y(x) := 1\!\!1_{\{|x|\leq 1\}} + \frac{1}{|x|} 1\!\!1_{\{|x|>1\}}.$$

In order to construct the Lévy martingale measure Q we need β so that β and Y satisfy the martingale condition. We have

$$\begin{aligned}
&\int_{\mathbb{R}^d} |xY(x) - h(x)|\, K(dx) \\
&\leq \int_{\mathbb{R}^d} |xY(x) - h_0(x)|\, K(dx) + \int_{\mathbb{R}^d} |h_0 - h(x)|\, K(dx) \\
&= \int_{\{|x|\leq 1\}} |x - h_0(x)|\, K(dx) + \int_{\{|x|>1\}} K(dx) + \int_{\mathbb{R}^d} |h_0 - h(x)|\, K(dx)
\end{aligned}$$

which is finite by Lemma C.1 and the fact that K integrates $|x|^2 \wedge 1$. So we can define β by

$$\beta := -c^{-1} \left(b + \int_{\mathbb{R}^d} (xY(x) - h(x))\, K(dx) \right).$$

Note that $f\left(\frac{1}{|x|}\right) = 1 - \frac{1+\log|x|}{|x|} < 1$ for $|x| > 1$, so that

$$\int_{\mathbb{R}^d} f(Y(x))\, K(dx) = \int_{\{|x|>1\}} f\left(\frac{1}{|x|}\right)\, K(dx) \leq K(\{|x| > 1\}) < \infty.$$

So by Corollary 2.17 there exists (on the path space or for a finite time horizon) a measure $\bar{Q}$ with finite relative entropy, Girsanov quantities β and Y and density process $Z = \mathcal{E}(N)$ with $N = \beta L^c + (Y - 1) * (\mu^L - \nu^P)$ under which UL is a Lévy process. Note that β was chosen such that Q is a martingale measure by Proposition 3.7. $\diamond$

3.3 Preservation of the Lévy Property

We now come to the main result of this chapter. We have seen how to reduce the relative entropy of a measure Q by choosing a "better" measure Q^ℓ which preserves the P-Lévy property of L. In Theorem 3.4 Q^ℓ was constructed in such a way that the Q^ℓ-characteristics are the "average" Q-characteristics of L so that UL ought to remain a (local) martingale under Q^ℓ if it is a local martingale under Q. The natural approach to show that the martingale condition holds for the new measure Q^ℓ is to interchange integration with respect to P with the integration with respect to $K(dx)ds$ as seen in the proof of the next proposition. However this involves Fubini's theorem which requires that the functions under consideration are integrable with respect to the product measure $P\otimes K\otimes\lambda_{[0,T]}$, and there is no reason why this should be the case in general. So we can prove that the martingale condition for Q^ℓ holds only if the big jumps of UL are Q-integrable. To make this precise, recall that by Lemma 3.5 $|U(xY-h)|*\nu_t^P$ is Q-integrable if and only if $|Ux-h(Ux)|Y*\nu_t^P$ is Q-integrable, and let us define a new set of martingale measures.

Definition 3.11 Let X be a semimartingale. We set

$$\mathcal{Q}_{\text{int}}^U(X) := \left\{ Q\in\mathcal{Q}_a^U(X) \,\middle|\, E_Q\left[|Ux-h(Ux)|*\nu_t^Q\right]<\infty \text{ for all } t\geq 0\right\},$$

where ν^Q is the Q-compensator of the jump measure of X. ◇

Proposition 3.12 *Let $Q\in\mathcal{Q}_e^U\cap\mathcal{Q}_f^U\cap\mathcal{Q}_{\text{int}}^U$ with Girsanov quantities β and Y. Then the measure Q^ℓ constructed from Q in Theorem 3.4 b) is in $\mathcal{Q}_e^U\cap\mathcal{Q}_f^U\cap\mathcal{Q}_\ell^U$.*

Proof. Let β^ℓ, Y^ℓ be the Girsanov quantities of Q^ℓ. From Theorem 3.4 we know that $Q^\ell \overset{\text{loc}}{\sim} P$, that L is a Q^ℓ-Lévy process and that $I_T(Q^\ell|P)<\infty$ for some $T>0$. Since the Girsanov quantities of Q^ℓ are deterministic and independent of time, the entropy process is a linear function in t, so $I_t(Q^\ell|P)<\infty$ for all $t\geq 0$. It remains to be shown that UL is a local Q^ℓ-martingale. By Proposition 3.7 we need to show that $\int_{\mathbb{R}^d}\left|U\left(xY^\ell(x)-h(x)\right)\right|\, K(dx)<\infty$ and that the martingale condition for UL is satisfied by β^ℓ and Y^ℓ.

Note that $E_Q\left[|U(xY(s,x)-h(x))|*\nu_T^P\right]<\infty$ for $Q\in\mathcal{Q}_e^U\cap\mathcal{Q}_f^U\cap\mathcal{Q}_{\text{int}}^U$ by Lemma 3.5 b), so we get by Fubini's theorem and Jensen's inequality

$$\begin{aligned}
\int_{\mathbb{R}^d}\left|U\left(xY^\ell(x)-h(x)\right)\right|\, K(dx) &= \int_{\mathbb{R}^d}\left|U\left(xE_Q\left[\frac{1}{T}\int_0^T Y(s,x)\, ds\right]-h(x)\right)\right|\, K(dx)\\
&= \int_{\mathbb{R}^d}\left|E_Q\left[\frac{1}{T}\int_0^T U(xY(s,x)-h(x))\, ds\right]\right|\, K(dx)\\
&\leq \frac{1}{T}\, E_Q\left[|U(xY(s,x)-h(x))|*\nu_T^P\right]<\infty.
\end{aligned}$$

It remains to show that the martingale condition for UL holds for β^ℓ and Y^ℓ. By Fubini's theorem we have

$$\begin{aligned}
&U\left(b + c\beta^\ell + \int_{\mathbb{R}^d}\left(xY^\ell(x) - h(x)\right)\ K(dx)\right)\\
&= U\left(b + cE_Q\left[\frac{1}{T}\int_0^T \beta_s\ ds\right] + \int_{\mathbb{R}^d}\left(xE_Q\left[\frac{1}{T}\int_0^T Y(s,x)\ ds\right] - h(x)\right)\ K(dx)\right)\\
&= \frac{1}{T}E_Q\left[\int_0^T U\left(b + c\beta_s + \int_{\mathbb{R}^d}(xY(s,x) - h(x))\ K(dx)\right)\ ds\right]\\
&= 0
\end{aligned}$$

since β and Y satisfy the martingale condition for UL. □

If Q is not in $\mathcal{Q}^U_{\text{int}}$, we do not know if the measure Q^ℓ as constructed in Theorem 3.4 preserves the local martingale property of UL if UL is a local Q-martingale. But nevertheless we should like to show that the P-Lévy property is preserved under the minimal entropy martingale measure for UL. The key for this is Proposition 3.16 which shows that $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f \cap \mathcal{Q}^U_{\text{int}}$ is "dense" in $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f$ in the sense of convergence in entropy. The idea of the overall proof is then as follows. We first approximate a local martingale measure Q (not necessarily in $\mathcal{Q}^U_{\text{int}}$) in the sense of convergence in entropy by a sequence $(Q^n)_{n\in\mathbb{N}}$ in $\mathcal{Q}^U_{\text{int}}$. Then we show in Proposition 3.17 that the corresponding measures $Q^{n,\ell}$ from Theorem 3.4 converge in entropy to Q^ℓ. If we then suppose that the entropy-minimizing martingale measure Q^E is *not* a Lévy martingale measure, then $I_t(Q^{E,\ell}|P) < I_t(Q^E|P)$ for some $t \geq 0$, so for n sufficiently large, the Lévy martingale measure $(Q^E)^{n,\ell}$ has smaller relative entropy than Q^E, a contradiction.

This procedure relates to Kabanov and Stricker (2001) who show that the set of local martingale measures with bounded densities is dense in $\mathcal{Q}_e$ in the sense of convergence in variation by a similar construction. Note, however, that convergence in variation of Q^n to Q in general does not imply convergence of the relative entropies.

The first lemmas show how to find the sequence $(Q^n)_{n\in\mathbb{N}}$ in $\mathcal{Q}^U_{\text{int}}$. The idea here is that $Q \in \mathcal{Q}^U_e$ is "locally" in $\mathcal{Q}^U_e \cap \mathcal{Q}^U_{\text{int}}$ in the sense that $|Ux - h(Ux)| * \nu^Q$ is locally Q-integrable by Lemma 3.6. So for $n \in \mathbb{N}$ we let $Q^n = Q$ on $\mathcal{F}_{\tau_n}$, where $(\tau_n)_{n\in\mathbb{N}}$ is the localizing sequence for the Q-integrability of $|Ux - h(Ux)| * \nu^Q$, and $Q^n = \bar{Q}$ "after" τ_n for some $\bar{Q} \in \mathcal{Q}^U_\ell$. Note that $\mathcal{Q}^U_\ell \subseteq \mathcal{Q}^U_{\text{int}}$; in fact $|Ux - h(Ux)| * \nu^{\bar{Q}}$ is deterministic and linear in time for $\bar{Q} \in \mathcal{Q}^U_\ell$, so that local integrability implies integrability in this case.

We start with the following lemma which states that the set of strictly positive martingales is stable under concatenation, see also Kabanov and Stricker (2002), Section 4.

Lemma 3.13 *Let Z and $\bar{Z}$ be two strictly positive martingales and τ a stopping time with $\tau < \infty$ P-a.s. Define*

$$\hat{Z}_t = Z_t \mathbb{1}_{[\![0,\tau]\!]} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}_t \mathbb{1}_{]\!]\tau,\infty[\![}.$$

Then $\hat{Z}$ is a strictly positive martingale.

Proof. Positivity of $\hat{Z}$ is immediate. We show $E[\hat{Z}_\sigma] = E[\hat{Z}_0] = 1$ for every bounded stopping time σ. So let σ be bounded by some $T < \infty$. Then Z^T and $\bar{Z}^T$ are uniformly integrable martingales, and we have by the optional sampling theorem $E[X^T_\sigma | \mathcal{F}_\tau] = X^T_{\sigma\wedge\tau}$ for $X \in \{Z, \bar{Z}\}$. Note further that $X^T_\sigma = X_\sigma$ for $X \in \{Z, \bar{Z}\}$ and that $\{\sigma \le \tau\}$ is $\mathcal{F}_\tau$-measurable. Thus we get

$$\begin{aligned}
E\big[\hat{Z}_\sigma\big] &= E\left[Z_\sigma 1\!\!1_{\{\sigma\le\tau\}} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}_\sigma 1\!\!1_{\{\sigma>\tau\}}\right] \\
&= E\left[Z^T_\sigma 1\!\!1_{\{\sigma\le\tau\}} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}^T_\sigma 1\!\!1_{\{\sigma>\tau\}}\right] \\
&= E\left[E\left[Z^T_\sigma|\mathcal{F}_\tau\right] 1\!\!1_{\{\sigma\le\tau\}} + \frac{Z_\tau}{\bar{Z}_\tau}E\left[\bar{Z}^T_\sigma|\mathcal{F}_\tau\right] 1\!\!1_{\{\sigma>\tau\}}\right] \\
&= E\left[Z^T_{\sigma\wedge\tau} 1\!\!1_{\{\sigma\le\tau\}} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}^T_{\sigma\wedge\tau} 1\!\!1_{\{\sigma>\tau\}}\right] \\
&= E\left[Z^T_\sigma 1\!\!1_{\{\sigma\le\tau\}} + Z^T_\tau 1\!\!1_{\{\sigma>\tau\}}\right] \\
&= E\left[Z^T_{\sigma\wedge\tau}\right] \\
&= 1,
\end{aligned}$$

since Z^T is a uniformly integrable martingale. □

Next we show how the concatenation of two stochastic exponentials carries over to their exponents. This will be helpful to construct a "concatenation" of probability measures for given Girsanov quantities.

Lemma 3.14 *Let $Z, \bar{Z}, \tau$ and $\hat{Z}$ be as above and suppose $Z = \mathcal{E}(N), \bar{Z} = \mathcal{E}(\bar{N})$ for some local martingales $N, \bar{N}$. Then $\hat{Z} = \mathcal{E}(\hat{N})$, where*

$$\hat{N}_t := N^\tau_t + \bar{N}_t - \bar{N}^\tau_t.$$

Proof. We show that $\hat{Z} = 1 + \int \hat{Z}_-\, d\hat{N}$. Note that $\hat{Z}_{t-} = Z_{t-}1\!\!1_{[\![0,\tau]\!]} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}_{t-}1\!\!1_{]\!]\tau,\infty[\![}$. We have

$$\begin{aligned}
\int_0^t \hat{Z}_{s-}\, d\hat{N}_s &= \int_0^t Z_{s-}1\!\!1_{[\![0,\tau]\!]}\, d\hat{N}_s + \frac{Z_\tau}{\bar{Z}_\tau}\int_0^t \bar{Z}_{s-}1\!\!1_{]\!]\tau,\infty[\![}\, d\hat{N}_s \\
&= \int_0^t Z_{s-}\, d\hat{N}^\tau_s + \frac{Z_\tau}{\bar{Z}_\tau}\left(\int_0^t \bar{Z}_{s-}\, d\hat{N}_s - \int_0^t \bar{Z}_{s-}\, d\hat{N}^\tau_s\right).
\end{aligned}$$

Now $\hat{N}^\tau = N^\tau$, so we get for the first term

$$\int_0^t Z_{s-}\, d\hat{N}^\tau_s = \int_0^t Z_{s-}\, dN^\tau_s = \left(\int Z_-\, dN\right)^\tau_t = Z_{t\wedge\tau} - 1 = Z_t 1\!\!1_{[\![0,\tau]\!]} + Z_\tau 1\!\!1_{]\!]\tau,\infty[\![} - 1,$$

whereas in the second term linearity of the stochastic integral in the integrator argument yields

$$\begin{aligned}\int_0^t \bar{Z}_{s-}\, d\hat{N}_s - \int_0^t \bar{Z}_{s-}\, d\hat{N}_s^\tau &= \int_0^t \bar{Z}_{s-}\, dN_s^\tau + \int_0^t \bar{Z}_{s-}\, d\bar{N}_s - \int_0^t \bar{Z}_{s-}\, d\bar{N}_s^\tau - \int_0^t \bar{Z}_{s-}\, dN_s^\tau \\ &= (\bar{Z}_t - 1) - (\bar{Z}_{t\wedge\tau} - 1) \\ &= \bar{Z}_t - \bar{Z}_t \mathbb{1}_{[\![0,\tau]\!]} - \bar{Z}_\tau \mathbb{1}_{]\!]\tau,\infty[\![}.\end{aligned}$$

So altogether we have

$$\begin{aligned}\int_0^t \hat{Z}_{s-}\, d\hat{N}_s &= Z_t \mathbb{1}_{[\![0,\tau]\!]} + Z_\tau \mathbb{1}_{]\!]\tau,\infty[\![} - 1 + \frac{Z_\tau}{\bar{Z}_\tau}\left(\bar{Z}_t - \bar{Z}_t \mathbb{1}_{[\![0,\tau]\!]} - \bar{Z}_\tau \mathbb{1}_{]\!]\tau,\infty[\![}\right) \\ &= Z_t \mathbb{1}_{[\![0,\tau]\!]} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}_t \mathbb{1}_{]\!]\tau,\infty[\![} - 1 \\ &= \hat{Z}_t - 1,\end{aligned}$$

hence the claim. □

The following arguments use the results from Section 2.3 concerning the construction of locally equivalent measures from given Girsanov quantities. In the case of an infinite time horizon the results can be stated only for the path space, so let $(\Omega, \mathcal{F}) = (\mathbb{D}, \mathcal{B}(\mathbb{D}))$ and P be a measure on $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ such that the coordinate process on $\mathbb{D}$, denoted by L, is a P-Lévy process. All results can also be stated for a finite time horizon $[0,T]$. In this case $(\Omega, \mathcal{F}, P)$ may be arbitrary and L is a P-Lévy process on $[0,T]$. Note that in both cases we assume $\mathbb{F} = \mathbb{F}^L(P)$ to be the P-augmentation of the filtration generated by L.

Lemma 3.15 *Let $Q, \bar{Q} \overset{\text{loc}}{\sim} P$ with Girsanov quantities (β, Y) and $(\bar{\beta}, \bar{Y})$, respectively. Then for every P-a.s. finite stopping time τ there exists a measure $\hat{Q} \overset{\text{loc}}{\sim} P$ with Girsanov quantities*

$$\begin{aligned}\hat{\beta} &= \beta \mathbb{1}_{[\![0,\tau]\!]} + \bar{\beta} \mathbb{1}_{]\!]\tau,\infty[\![}, \\ \hat{Y} &= Y \mathbb{1}_{[\![0,\tau]\!]} + \bar{Y} \mathbb{1}_{]\!]\tau,\infty[\![}.\end{aligned}$$

Proof. Let Z and $\bar{Z}$ be the density processes of Q and $\bar{Q}$. Then Lemma 3.13 yields that $\hat{Z} = Z\mathbb{1}_{[\![0,\tau]\!]} + \frac{Z_\tau}{\bar{Z}_\tau}\bar{Z}\mathbb{1}_{]\!]\tau,\infty[\![}$ is a P-martingale, which is strictly positive if Z and $\bar{Z}$ are. So by Theorem 2.10 there exists $\hat{Q} \overset{\text{loc}}{\sim} P$ with density process $\hat{Z}$. From Corollary 2.9 we know that Z and $\bar{Z}$ are given by $Z = \mathcal{E}(N)$ and $\bar{Z} = \mathcal{E}(\bar{N})$ with $N = \int \beta dL^c + (Y-1) * (\mu^L - \nu^P)$ and $\bar{N} = \int \bar{\beta}\, dL^c + (\bar{Y} - 1) * (\mu^L - \nu^P)$, so by Lemma 3.14 we know that $\hat{Z} = \mathcal{E}(\hat{N})$ with $\hat{N} = N^\tau + \bar{N} - \bar{N}^\tau$.

We define the quantities β' and Y' by

$$\begin{aligned}\beta'_s(\omega) &= \beta_s(\omega)\mathbb{1}_{[\![0,\tau]\!]}(\omega,s) + \bar{\beta}_s(\omega)\mathbb{1}_{]\!]\tau,\infty[\![}(\omega,s), \\ Y'(\omega;s,x) &= Y(\omega;s,x)\mathbb{1}_{[\![0,\tau]\!]}(\omega,s) + \bar{Y}(\omega;s,x)\mathbb{1}_{]\!]\tau,\infty[\![}(\omega,s).\end{aligned}$$

Since Y and $\bar{Y}$ are predictable functions and since $1\!\!1_{[\![0,\tau]\!]}$ is predictable, Y' is a predictable function, and with the same argument β' is a predictable process. Furthermore,

$$\begin{aligned}\left(1-\sqrt{Y'}\right)^2 * \nu_t^P &= \left(1-\sqrt{Y}\right)^2 1\!\!1_{[\![0,\tau]\!]} * \nu_t^P + \left(1-\sqrt{\bar{Y}}\right)^2 1\!\!1_{]\!]\tau,\infty[\![} * \nu_t^P \\ &\leq \left(1-\sqrt{Y}\right)^2 * \nu_t^P + \left(1-\sqrt{\bar{Y}}\right)^2 * \nu_t^P\end{aligned}$$

which is locally P-integrable since $Y-1$ and $\bar{Y}-1$ are integrable with respect to $\mu^L - \nu^P$. So by Theorem 1.22 c) $Y'-1$ is integrable with respect to $\mu^L-\nu^P$. We define the process N' by

$$\begin{aligned}N'_t &= \int_0^t \beta'_s \, dL_s^c + (Y'-1) * (\mu^L - \nu^P)_t \\ &= \int_0^t \beta_s 1\!\!1_{[\![0,\tau]\!]} \, dL_s^c + (Y-1)1\!\!1_{[\![0,\tau]\!]} * (\mu^L-\nu^P)_t + \int_0^t \bar{\beta}_s \, dL_s^c + (\bar{Y}-1) * (\mu^L - \nu^P)_t \\ &\quad - \int_0^t \bar{\beta}_s 1\!\!1_{[\![0,\tau]\!]} \, dL_s^c - (\bar{Y}-1)1\!\!1_{[\![0,\tau]\!]} * (\mu^L-\nu^P)_t \\ &= N_t^\tau + \bar{N}_t - \bar{N}_t^\tau \\ &= \hat{N}_t,\end{aligned}$$

and we see that $\hat{Z} = \mathcal{E}(N')$, so β' and Y' are the Girsanov quantities of $\hat{Q}$ by Proposition 2.16. □

We are now in a position to show that $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_{\text{int}}^U$ is "dense" in $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U$ in the sense of convergence in entropy. Recall, however, that convergence in entropy is not compatible with convergence of probability measures in a classical way in that it does not induce a topology on $\{Q \overset{\text{loc}}{\ll} P\}$, so we put "dense" in quotation marks.

Proposition 3.16 *Let $Q \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U$ and suppose $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U \neq \emptyset$. Then there exists a sequence $(Q^n)_{n\in\mathbb{N}}$ in $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_{\text{int}}^U$ with $I_t(Q^n|P) \overset{n\to\infty}{\longrightarrow} I_t(Q|P)$ for all $t \geq 0$.*

Proof. Let β and Y be the Girsanov quantities of Q. By assumption UL is a local Q-martingale, so $|Ux - h(Ux)|Y * \nu^P$ is continuous and finite-valued by Proposition 3.7. Thus $|Ux-h(Ux)|Y*\nu^P$ is locally Q-integrable, so there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times with $\tau_n \uparrow \infty$ Q-a.s. such that $E_Q[|Ux-h(Ux)|Y * \nu^P_{\tau_n}] < \infty$ for all $n \in \mathbb{N}$. We construct a measure Q^n which coincides with Q on $\mathcal{F}_{\tau_n}$ and is a Lévy martingale measure for UL, denoted by $\bar{Q}$, "after" τ_n via Lemma 3.15.

Let $Z = \mathcal{E}(N)$ be the density process of Q with $N = \int \beta \, dL^c + (Y-1) * (\mu^L - \nu^P)$ and let $\bar{Q} \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$ be a Lévy martingale measure with Girsanov quantities $\bar{\beta}$ and $\bar{Y}$. This uses the assumption that $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U \neq \emptyset$. Then $\bar{Z} = \mathcal{E}(\bar{N})$ with $\bar{N} = \bar{\beta}L^c + (\bar{Y}-1)*(\mu^L-\nu^P)$ is the density process of $\bar{Q}$ with respect to P. So for all $n \in \mathbb{N}$, Lemma 3.15 gives us a measure $Q^n \overset{\text{loc}}{\sim} P$ with Girsanov quantities

$$\begin{aligned}\beta_s^n &= \beta_s 1\!\!1_{[\![0,\tau_n]\!]} + \bar{\beta} 1\!\!1_{]\!]\tau_n,\infty[\![}, \\ Y^n(s,x) &= Y(s,x) 1\!\!1_{[\![0,\tau_n]\!]} + \bar{Y}(x) 1\!\!1_{]\!]\tau_n,\infty[\![}.\end{aligned}$$

We next show that Q^n is a local martingale measure for UL which satisfies the integrability condition of $\mathcal{Q}^U_{\text{int}}$, i.e. $E_{Q^n}\left[|Ux - h(Ux)| * \nu_t^{Q^n}\right] = E_{Q^n}\left[|Ux - h(Ux)|Y^n * \nu_t^P\right] < \infty$ for $t \geq 0$. Since Q and $\bar{Q}$ are local martingale measures for UL, we see that

$$\begin{aligned}
|U(xY^n - h)| * \nu_t^P &= \left|Ux\left(Y(s,x)1\!\!1_{[\![0,\tau_n]\!]} + \bar{Y}(x)1\!\!1_{]\!]\tau_n,\infty[\![}\right) - Uh(x)\right| * \nu_t^P \\
&\leq |U(xY(s,x) - h(x))|1\!\!1_{[\![0,\tau_n]\!]} * \nu_t^P + |U(x\bar{Y}(x) - h(x))|1\!\!1_{]\!]\tau_n,\infty[\![} * \nu_t^P \\
&\leq |U(xY - h)| * \nu_t^P + |U(x\bar{Y} - h)| * \nu_t^P < \infty
\end{aligned}$$

Q^n-a.s. for all $t \geq 0$ by Proposition 3.7, and the martingale condition is obviously satisfied by the choice of β^n and Y^n, so that Q^n is a local martingale measure for UL. For the Q^n-integrability of $|Ux - h(Ux)|Y^n * \nu_t^P$, note that $Q^n = Q$ on $\mathcal{F}_{\tau_n}$ and that $\bar{Y}$ is deterministic and independent of time, so we get

$$\begin{aligned}
&E_{Q^n}\left[|Ux - h(Ux)|Y^n * \nu_t^P\right] \\
&= E_{Q^n}\left[|Ux - h(Ux)|Y1\!\!1_{[\![0,\tau_n]\!]} * \nu_t^P\right] + E_{Q^n}\left[|Ux - h(Ux)|\bar{Y}1\!\!1_{]\!]\tau_n,\infty]\!]} * \nu_t^P\right] \\
&= E_{Q^n}\left[|Ux - h(Ux)|Y1\!\!1_{[\![0,\tau_n]\!]} * \nu_t^P\right] + \int_{\mathbb{R}^d} |Ux - h(Ux)|\bar{Y}(x)\ K(dx)\ E_{Q^n}\left[\int_0^t 1\!\!1_{]\!]\tau_n,\infty[\![}\ ds\right] \\
&\leq E_Q\left[|Ux - h(Ux)|Y * \nu_{\tau_n}^P\right] + t\int_{\mathbb{R}^d} |Ux - h(Ux)|\bar{Y}(x)\ K(dx),
\end{aligned}$$

since $\int_0^t 1\!\!1_{]\!]\tau_n,\infty[\![}\ ds \leq t$. Now the first term is finite by the choice of τ_n and the second by Proposition 3.7 and Lemma 3.5 a).

We finally show that the entropy process of Q with respect to P is finite-valued and that Q^n converges in entropy to Q. From Lemma 3.3 b) we know that

$$I_t(Q^n|P) = \frac{1}{2}E_{Q^n}\left[\int_0^t (\beta_s^n)^{\text{tr}} c\beta_s^n\ ds\right] + E_{Q^n}\left[f(Y^n) * \nu_t^P\right]$$

and

$$I_t(Q|P) = \frac{1}{2}E_Q\left[\int_0^t (\beta_s)^{\text{tr}} c\beta_s\ ds\right] + E_Q\left[f(Y) * \nu_t^P\right].$$

So again using the fact that $Q^n = Q$ on $\mathcal{F}_{\tau_n}$ and that τ_n is $\mathcal{F}_{\tau_n}$-measurable, we see that

$$\begin{aligned}
E_{Q^n}\left[\int_0^t (\beta_s^n)^{\text{tr}} c\beta_s^n\ ds\right] &= E_{Q^n}\left[\int_0^t (\beta_s)^{\text{tr}} c\beta_s 1\!\!1_{[\![0,\tau_n]\!]}(s)\ ds\right] + E_{Q^n}\left[\int_0^t \bar{\beta}^{\text{tr}} c\bar{\beta}1\!\!1_{]\!]\tau_n,\infty[\![}(s)\ ds\right] \\
&= E_Q\left[\int_0^t (\beta_s)^{\text{tr}} c\beta_s 1\!\!1_{[\![0,\tau_n]\!]}(s)\ ds\right] + \bar{\beta}^{\text{tr}} c\bar{\beta}\ E_Q[(t - \tau_n) \vee 0].
\end{aligned}$$

Now $\tau_n \uparrow \infty$ Q-a.s., so $\int_0^t (\beta_s)^{\text{tr}} c\beta_s 1\!\!1_{[\![0,\tau_n]\!]}\ ds \uparrow \int_0^t (\beta_s)^{\text{tr}} c\beta_s\ ds$ and $(t - \tau_n) \vee 0 \downarrow 0$ Q-a.s., and we get by monotone convergence

$$E_Q\left[\int_0^t (\beta_s)^{\text{tr}} c\beta_s 1\!\!1_{[\![0,\tau_n]\!]}\ ds\right] + \bar{\beta}^{\text{tr}} c\bar{\beta}\ E_Q[(t - \tau_n) \vee 0] \to E_Q\left[\int_0^t (\beta_s)^{\text{tr}} c\beta_s\ ds\right]$$

as $n \to \infty$. Similarly we have

$$\begin{aligned}
&E_{Q^n}\left[f(Y^n) * \nu_t^P\right] \\
&= E_{Q^n}\left[f(Y \mathbb{1}_{[\![0,\tau_n]\!]} + \bar{Y}\mathbb{1}_{]\!]\tau_n,\infty[\![}) * \nu_t^P\right] \\
&= E_{Q^n}\left[f(Y)\mathbb{1}_{[\![0,\tau_n]\!]} * \nu_t^P\right] + E_{Q^n}\left[f(\bar{Y})\mathbb{1}_{]\!]\tau_n,\infty[\![} * \nu_t^P\right] \\
&= E_Q\left[\int_0^t \int_{\mathbb{R}^d} f(Y(s,x))\mathbb{1}_{[\![0,\tau_n]\!]}(s)\ K(dx)ds\right] + \int_{\mathbb{R}^d} f(\bar{Y}(x))\ K(dx)\ E_Q\left[(t-\tau_n)\vee 0\right] \\
&\overset{n\to\infty}{\longrightarrow} E_Q\left[\int_0^t \int_{\mathbb{R}^d} f(Y(s,x))\ K(dx)ds\right] \\
&= E_Q\left[f(Y(s,x)) * \nu_t^P\right]
\end{aligned}$$

by monotone convergence. This finishes the proof. □

Proposition 3.17 *Let $Q \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U$ and suppose $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U \neq \emptyset$. Choose the sequence $(Q^n)_{n\in\mathbb{N}}$ in $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_{\mathrm{int}}^U$ as in the proof of Proposition 3.16 and denote for some $T > 0$ the corresponding measures from Theorem 3.4 b) by Q^ℓ and $Q^{n,\ell} = (Q^n)^\ell$, respectively. Then $I_t(Q^{n,\ell}|P) \to I_t(Q^\ell|P)$ as $n \to \infty$ for all $t \geq 0$.*

Proof. Recall that the Girsanov quantities of Q^n are given by

$$\begin{aligned}
\beta_s^n &= \beta_s \mathbb{1}_{[\![0,\tau_n]\!]} + \bar{\beta}\mathbb{1}_{]\!]\tau_n,\infty[\![}, \\
Y^n(s,x) &= Y(s,x)\mathbb{1}_{[\![0,\tau_n]\!]} + \bar{Y}(x)\mathbb{1}_{]\!]\tau_n,\infty[\![},
\end{aligned}$$

where (β, Y) are the Girsanov quantities of Q and $(\bar{\beta}, \bar{Y})$ are the Girsanov quantities of some $\bar{Q} \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$. Fix some $T > 0$. Then by the construction in Theorem 3.4 we get the Girsanov quantities of $Q^{n,\ell}$ as

$$\beta^{n,\ell} = E_{Q^n}\left[\frac{1}{T}\int_0^T \beta_s^n\ ds\right] = E_Q\left[\frac{1}{T}\int_0^T \beta_s \mathbb{1}_{[\![0,\tau_n]\!]}(s)\ ds\right] + \frac{1}{T}\bar{\beta}\ E_Q[(T-\tau_n)\vee 0]$$

and

$$\begin{aligned}
Y^{n,\ell}(x) &= E_{Q^n}\left[\frac{1}{T}\int_0^T Y^n(s,x)\ ds\right] \\
&= E_Q\left[\frac{1}{T}\int_0^T Y(s,x)\mathbb{1}_{[\![0,\tau_n]\!]}(s)\ ds\right] + \bar{Y}(x)\ \frac{1}{T}\ E_Q\left[(T-\tau_n)\vee 0\right].
\end{aligned}$$

By Lemma 3.3 b) the entropy processes of $Q^{n,\ell}$ and Q^ℓ are given by

$$\begin{aligned}
I_t(Q^{n,\ell}|P) &= \left(\frac{1}{2}(\beta^{n,\ell})^{\mathrm{tr}} c \beta^{n,\ell} + \int_{\mathbb{R}^d} f(Y^{n,\ell}(x))\ K(dx)\right) t, \\
I_t(Q^\ell|P) &= \left(\frac{1}{2}(\beta^\ell)^{\mathrm{tr}} c \beta^\ell + \int_{\mathbb{R}^d} f(Y^\ell(x))\ K(dx)\right) t,
\end{aligned}$$

so if we show that $\beta^{n,\ell} \to \beta^\ell$ and $\int_{\mathbb{R}^d} f(Y^{n,\ell}(x))\ K(dx) \to \int_{\mathbb{R}^d} f(Y^\ell(x))\ K(dx)$ as $n \to \infty$, we immediately get convergence of the relative entropies.

Now since $\tau_n \uparrow \infty$ Q-a.s. and since $\int_0^T |\beta_s|\ ds$ is Q-integrable (and thus finite-valued) by Lemma 3.3 a), we get

$$\int_0^T \beta_s \mathbb{1}_{[\![0,\tau_n]\!]}(s)\ ds \overset{n\to\infty}{\longrightarrow} \int_0^T \beta_s\ ds \quad Q\text{-a.s.}$$

Furthermore, $\left|\int_0^T \beta_s \mathbb{1}_{[\![0,\tau_n]\!]}(s)\ ds\right| \leq \int_0^T |\beta_s|\ ds$ Q-a.s., and since $\int_0^T |\beta_s|\ ds$ is Q-integrable by Lemma 3.3 a), we get

$$E_Q\left[\frac{1}{T}\int_0^T \beta_s \mathbb{1}_{[\![0,\tau_n]\!]}(s)\ ds\right] \overset{n\to\infty}{\longrightarrow} E_Q\left[\frac{1}{T}\int_0^T \beta_s\ ds\right] = \beta^\ell$$

by dominated convergence. Finally $\tau_n \uparrow \infty$ implies $(T-\tau_n)\vee 0 \downarrow 0$, so $E_Q[(T-\tau_n)\vee 0] \overset{n\to\infty}{\longrightarrow} 0$ by monotone convergence, so that altogether we have $\beta^{n,\ell} \to \beta^\ell$.

Concerning convergence of $\int_{\mathbb{R}^d} f(Y^{n,\ell}(x))\ K(dx)$, we note that with the same arguments as above we have $Y^{n,\ell}(x) \overset{n\to\infty}{\longrightarrow} Y^\ell(x)$ for all $x \in \operatorname{supp} K$, and since f is continuous, we get $f(Y^{n,\ell}(x)) \overset{n\to\infty}{\longrightarrow} f(Y^\ell(x))$ for all $x \in \operatorname{supp} K$. To find a K-integrable dominating function for $f(Y^{n,\ell}(x))$, we use Jensen's inequality (recall that f is convex and nonnegative):

$$\begin{aligned}
f(Y^{n,\ell}) &= f\left(E_{Q^n}\left[\frac{1}{T}\int_0^T Y^n(s,x)\ ds\right]\right)\\
&\leq E_{Q^n}\left[\frac{1}{T}\int_0^T f(Y^n(s,x))\ ds\right]\\
&= E_{Q^n}\left[\frac{1}{T}\int_0^T f\Big(Y(s,x)\mathbb{1}_{[\![0,\tau_n]\!]}(s) + \bar{Y}(x)\mathbb{1}_{]\!]\tau_n,\infty[\![}(s)\Big)\ ds\right]\\
&= E_{Q^n}\left[\frac{1}{T}\int_0^T f(Y(s,x))\mathbb{1}_{[\![0,\tau_n]\!]}(s) + f(\bar{Y}(x))\mathbb{1}_{]\!]\tau_n,\infty[\![}(s)\ ds\right]\\
&= E_Q\left[\frac{1}{T}\int_0^T f(Y(s,x))\mathbb{1}_{[\![0,\tau_n]\!]}(s)\ ds\right] + \frac{1}{T}\ f(\bar{Y}(x))\ E_Q[(T-\tau_n)\vee 0]\\
&\leq E_Q\left[\frac{1}{T}\int_0^T f(Y(s,x))\ ds\right] + f(\bar{Y}(x)).
\end{aligned}$$

However, Q and $\bar{Q}$ have finite relative entropy, so from Lemma 3.3 a) we know that

$$E_Q\left[\int_{\mathbb{R}^d}\int_0^T f(Y(s,x))\ dsK(dx)\right] = E_Q\left[f(Y) * \nu_T^P\right] < \infty,$$

so we get K-integrability of $E_Q\left[\frac{1}{T}\int_0^T f(Y(s,x))\ ds\right]$ by Fubini's theorem, and with the same argument K-integrability of $f(\bar{Y}(x))$. So dominated convergence yields

$$\int_{\mathbb{R}^d} f(Y^{n,\ell}(x))\ K(dx) \overset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} f(Y^\ell(x))\ K(dx),$$

which finishes the proof. $\square$

We are now in a position to state the following sufficient criterion for the existence of Q^E. Recall that the entropy-minimizing Lévy martingale measure Q^E_ℓ for UL was defined to be the measure which minimizes the entropy process pointwise over all $Q \in \mathcal{Q}^U_\ell$, i.e.

$$I_t(Q^E_\ell|P) \leq I_t(Q|P) \text{ for all } Q \in \mathcal{Q}^U_\ell, t \geq 0.$$

Corollary 3.18 *Suppose $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f \cap \mathcal{Q}^U_\ell \neq \emptyset$. If Q^E_ℓ exists, then Q^E exists.*

Proof. Let Q^E_ℓ be the entropy-minimizing Lévy martingale measure and suppose that Q^E does not exist. Then there exist $T_0 \geq 0$ and $Q \in \mathcal{Q}^U_f$ with $I_{T_0}(Q|P) = I_{T_0}(Q^E_\ell|P) - \delta$ for some $\delta > 0$. Let $Q' \in \mathcal{Q}^U_e \cap \mathcal{Q}^U_f$ and define

$$Q^\varepsilon := (1-\varepsilon)Q + \varepsilon Q' \text{ for } 0 < \varepsilon < 1.$$

Then $Q^\varepsilon \in \mathcal{Q}^U_e \cap \mathcal{Q}^U_f$, and for ε sufficiently small we have

$$I_{T_0}(Q^\varepsilon|P) \leq (1-\varepsilon)I_{T_0}(Q|P) + \varepsilon I_{T_0}(Q'|P) < I_{T_0}(Q^E_\ell|P)$$

by Lemma 1.10. Let $Q^{\varepsilon,\ell} = (Q^\varepsilon)^\ell$ be the corresponding measure from Theorem 3.4 for $T = T_0$, so that $I_{T_0}(Q^{\varepsilon,\ell}|P) \leq I_{T_0}(Q^\varepsilon|P) < I_{T_0}(Q^E_\ell|P)$. From Proposition 3.17 we get a sequence $(Q^{n,\varepsilon,\ell})_{n\in\mathbb{N}}$ in $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f \cap \mathcal{Q}^U_\ell$ with $I_{T_0}(Q^{n,\varepsilon,\ell}|P) \to I_{T_0}(Q^{\varepsilon,\ell}|P)$ as $n \to \infty$. So for n sufficiently large we have $I_{T_0}(Q^{n,\varepsilon,\ell}|P) < I_{T_0}(Q^E_\ell|P)$, a contradiction to the optimality of Q^E_ℓ. □

Theorem 3.19 *Suppose $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f \cap \mathcal{Q}^U_\ell \neq \emptyset$. If Q^E exists, then L is a Lévy process under Q^E.*

Proof. Suppose L is not a Lévy process under Q^E, then there exists $T_0 > 0$ such that L is not a Q^E-Lévy process on $[0, T_0]$. Let $Q^{E,\ell} = (Q^E)^\ell$ be the measure obtained from Q^E by Theorem 3.4 b) for $T = T_0$, then Theorem 3.4 c) yields $I_{T_0}(Q^{E,\ell}|P) < I_{T_0}(Q^E|P)$. Note that this is not yet a contradiction to the optimality of Q^E since we do not know whether UL is a local martingale under $Q^{E,\ell}$.

Let $(Q^{E,n})_{n\in\mathbb{N}}$ in $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f \cap \mathcal{Q}^U_{\text{int}}$ be the sequence of martingale measures for UL from Proposition 3.16 with $I_t(Q^{E,n}|P) \stackrel{n\to\infty}{\longrightarrow} I_t(Q^E|P)$ for all $t \geq 0$ and let $Q^{E,n,\ell} = (Q^{E,n})^\ell$ be the corresponding (Lévy) martingale measures for UL obtained from $Q^{E,n}$ by Theorem 3.4. By Proposition 3.17 we know that

$$I_{T_0}(Q^{E,n,\ell}|P) \stackrel{n\to\infty}{\longrightarrow} I_{T_0}(Q^{E,\ell}|P) < I_{T_0}(Q^E|P).$$

So for n sufficiently large we have $I_{T_0}(Q^{E,n,\ell}|P) < I_{T_0}(Q^E|P)$, the desired contradiction. □

Corollary 3.20 *Suppose $\mathcal{Q}^U_e \cap \mathcal{Q}^U_f \cap \mathcal{Q}^U_\ell \neq \emptyset$. If Q^E_ℓ exists, then $Q^E = Q^E_\ell$ and $Q^E_\ell \stackrel{\text{loc}}{\sim} P$.*

Proof. By Corollary 3.18 we know that Q^E exists, and by definition Q^E is optimal in $\mathcal{Q}_f^U$. By Theorem 3.19 $Q^E \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$, so Q^E is optimal in $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U \subseteq \mathcal{Q}_f$, hence $Q^E = Q_\ell^E$. The local equivalence of Q_ℓ^E and P then immediately follows from $Q^E \overset{\text{loc}}{\sim} P$ (cf. Lemma 1.15). □

Remark 3.21 Theorem 3.19 implies that in order to determine Q^E it suffices to find a martingale measure which is optimal in $\mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$, whereas Corollary 3.20 shows that this measure must be locally equivalent to P. So altogether we have to look for the optimal measure in $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$, and thus the assumption $\mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U \neq \emptyset$ is a natural condition to make and does not imply a loss of generality.

However notice that we still lack sufficient conditions on L, respectively the P-Lévy characteristics of L, which ensure the existence of the entropy- minimizing martingale measure. In the next chapter we introduce natural conditions under which Q^E exists, and where we can even specify the density of Q^E with respect to P, see Theorem 4.7. ◇

Chapter 4

Applications

In this chapter we introduce several models where the (discounted) stock price, denoted by X, is the solution of a stochastic differential equation driven by a P-Lévy process L. In many cases X is a local Q-martingale if and only if L is a local Q-martingale, so $Q^E(X)$, the entropy-minimizing martingale measure for X, coincides with $Q^E(L)$, the entropy-minimizing martingale measure for L.

In Section 4.1 we start with an informal argument why the entropy-minimizing martingale measure, provided it exists, should coincide with the so called Esscher martingale measure. Esscher measures and the rôle they play in our context are investigated in Section 4.2. Finally in Sections 4.3 and 4.4 we analyze a generalization of the Black-Scholes model, where the driving processes are Lévy processes (rather than only Brownian motion), and a model with stochastic volatility. There we give conditions under which the entropy-minimizing martingale measure exists and we specify its density process.

4.1 Identification of the Optimal Measure

Let L be a P-Lévy process with P-Lévy characteristics (b, c, K). In the following we investigate martingale measures for L, so the matrix U introduced in Chapter 3 is the identity matrix, and in the sequel we suppress the dependence of $\mathcal{Q}_x$ on L and U. Recall that in Chapter 3 we have seen that $Q^E(L) \in \mathcal{Q}_e(L) \cap \mathcal{Q}_f(L) \cap \mathcal{Q}_\ell(L)$, so the problem of finding Q^E reduces to a deterministic variational problem, and we start with the following, admittedly informal, argument. Since any candidate for Q^E must be in $\mathcal{Q}_e \cap \mathcal{Q}_f \cap \mathcal{Q}_\ell$, let $Q \overset{\text{loc}}{\sim} P$ be some locally equivalent local martingale measure for L with finite relative entropy and deterministic Girsanov quantities β and $Y = Y(x)$. Then by Lemma 3.3 b) the entropy process is given by

$$I_t(Q|P) = \left(\frac{1}{2}\beta^{\text{tr}} c\beta + \int_{\mathbb{R}^d} f(Y(x))\ K(dx)\right) t,$$

and by the martingale condition, Y and β are related by

$$c\beta = -b - \int_{\mathbb{R}^d} (xY(x) - h(x)) \ K(dx).$$

Suppose that c is regular, then solving for β and plugging in yields

$$\beta^{\mathrm{tr}} c\beta = (b + k(Y))^{\mathrm{tr}} \, c^{-1} \, (b + k(Y)) \, ,$$

where $k(Y) := \int_{\mathbb{R}^d} (xY(x) - h(x)) \ K(dx)$, so that

$$I(Q|P)_t = \left(\frac{1}{2} \, (b + k(Y))^{\mathrm{tr}} \, c^{-1} \, (b + k(Y)) + \int_{\mathbb{R}^d} f(Y(x)) \ K(dx) \right) t =: \bar{I}(Y) \ t.$$

Now we want to minimize $\bar{I}(Y)$ over functions $Y = Y(x)$. If some Y^* is optimal, we have for all Y and all $\varepsilon > 0$

$$\bar{I} \left((1 - \varepsilon) Y^* + \varepsilon Y \right) \geq \bar{I}(Y^*),$$

and thus

$$\begin{aligned}
0 \ \leq \ & \bar{I}(Y^* + \varepsilon(Y - Y^*)) - \bar{I}(Y^*) \\
= \ & \frac{1}{2} \left(b + \int (xY^* + \varepsilon x (Y - Y^*) - h) \ dK \right)^{\mathrm{tr}} c^{-1} \left(b + \int (xY^* - h) \ dK \right) + \\
& + \frac{1}{2} \, \varepsilon \left(b + \int (xY^* - h) \ dK \right)^{\mathrm{tr}} c^{-1} \left(\int x(Y - Y^*) \ dK \right) + \\
& + \frac{1}{2} \, \varepsilon^2 \left(\int x(Y - Y^*) \ dK \right)^{\mathrm{tr}} c^{-1} \left(\int x(Y - Y^*) \ dK \right) - \\
& - \frac{1}{2} \left(b + \int (xY^* - h) \ dK \right)^{\mathrm{tr}} c^{-1} \left(b + \int (xY^* - h) \ dK \right) + \\
& + \int (f(Y^* + \varepsilon(Y - Y^*)) - f(Y^*)) \ dK \\
= \ & \varepsilon \left(\int x(Y - Y^*) \ dK \right)^{\mathrm{tr}} c^{-1} \left(b + \int (xY^* - h) \ dK \right) + \\
& + \frac{1}{2} \, \varepsilon^2 \left(\int x(Y - Y^*) \ dK \right)^{\mathrm{tr}} c^{-1} \left(\int x(Y - Y^*) \ dK \right) + \\
& + \int (f(Y^* + \varepsilon(Y - Y^*)) - f(Y^*)) \ dK.
\end{aligned}$$

Now dividing by ε and taking the limit as ε tends to 0, we obtain

$$0 \leq \left(\int x(Y - Y^*) \ dK \right)^{\mathrm{tr}} c^{-1} \left(b + \int (xY^* - h) \ dK \right) + \int (Y - Y^*) f'(Y^*) \ dK.$$

If we then set $Y = (1 \pm \delta) Y^*$ for $\delta > 0$, this leads to

$$0 \leq \pm \, \delta \left(\left(\int xY^* \ dK \right)^{\mathrm{tr}} c^{-1} \left(b + \int (xY^* - h) \ dK \right) + \int Y^* f'(Y^*) \ dK \right),$$

which is only possible if

$$\begin{aligned} 0 &= \left(\int xY^* \, dK\right)^{\mathrm{tr}} c^{-1} \left(b + \int (xY^* - h) \, dK\right) + \int Y^* f'(Y^*) \, dK \\ &= \int \left(-(\beta^*)^{\mathrm{tr}} x + \log Y^*\right) Y^* \, dK, \end{aligned}$$

where $\beta^* = \beta^*(Y^*) = -c^{-1}\left(b + \int (xY^* - h)\, dK\right)$ is the optimal β from the martingale condition; note that $f'(y) = \log y$.

So morally we need

$$\log Y^*(x) - (\beta^*)^{\mathrm{tr}} x = 0,$$

or

$$Y^*(x) = e^{(\beta^*)^{\mathrm{tr}} x}.$$

Hence the optimal measure Q^* should admit Girsanov quantities $\beta^* = u^*$ and $Y^*(x) = e^{(u^*)^{\mathrm{tr}} x}$ for some $u^* \in \mathbb{R}^d$. The martingale condition then reads

$$b + cu^* + \int \left(xe^{(u^*)^{\mathrm{tr}} x} - h(x)\right) \, K(dx) = 0.$$

Note that in the above argument we have not shown existence of a measure with these Girsanov quantities. However the special form of $Y^*(x) = e^{(\beta^*)^{\mathrm{tr}} x}$ is satisfied by the Esscher measures. In addition, an Esscher measure is a good candidate for Q^E since it preserves the Lévy property of L. This will be seen in the following section.

4.2 Esscher Measures

In this section we define Esscher measures for Lévy processes, derive some useful properties of Esscher measures needed in our context and calculate their Girsanov quantities. The definition of Esscher measures used here is taken from Shiryaev (1999), VII.3c.

Let L be a P-Lévy process with P-Lévy characteristics (b, c, K) and fix a $d \times d$-matrix U. Recall Theorem A.7 for the following statements: For $u \in \mathbb{R}^d$ with $\int_{\{|x|>1\}} e^{u^{\mathrm{tr}} x} \, K(dx) < \infty$ the function Ψ, defined by

$$\Psi(u) = b^{\mathrm{tr}} u + \frac{1}{2} u^{\mathrm{tr}} c u + \int_{\mathbb{R}^d} \left(e^{u^{\mathrm{tr}} x} - 1 - (u^{\mathrm{tr}} x)\mathbb{1}_{\{|x| \le 1\}}\right) \, K(dx),$$

is finite-valued, and

$$E_P[\exp(u^{\mathrm{tr}} L_t)] = \exp(t \; \Psi(u)).$$

Definition 4.1 Let $u \in \mathbb{R}^d$ with $E_P[\exp(u^{\mathrm{tr}} L_1)] < \infty$ and define

$$Z_t^u = \frac{\exp(u^{\mathrm{tr}} L_t)}{E_P[\exp(u^{\mathrm{tr}} L_t)]}.$$

It is straightforward, using the Lévy structure of L under P, to see that Z^u is a strictly positive P-martingale. So Z^u is the density process of a measure $Q^u \overset{\text{loc}}{\sim} P$ on $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ by Theorem 2.10. Any such Q^u is called *Esscher measure for L* or simply Esscher measure if there is no ambiguity about L. If Q^u is in addition a martingale measure for UL, we call Q^u *Esscher martingale measure for UL.* $\diamond$

The next proposition is cited from Shiryaev (1999), Theorem VII.3.1.

Proposition 4.2 *Let $u \in \mathbb{R}^d$ with $E_P[\exp(u^{\text{tr}} L_1)] < \infty$ and let Q^u be an Esscher measure. Then L is a Lévy process under Q^u.*

Proof. Let Z^u be the density process of Q^u with respect to P. By Theorem A.7 we know $E_P[\exp(u^{\text{tr}} L_t)] = \exp(t\ \Psi(u))$, so that

$$Z_t^u = \exp\left(u^{\text{tr}} L_t - t\ \Psi(u)\right).$$

With this, we see that for $w \in \mathbb{R}^d$ and $0 \leq s \leq t$

$$\begin{aligned}
E_{Q^u}[\exp(iw^{\text{tr}}(L_t - L_s))|\mathcal{F}_s] &= E_P\left[\frac{Z_t^u}{Z_s^u}\ \exp(i\ w^{\text{tr}}(L_t - L_s))\middle|\ \mathcal{F}_s\right] \\
&= E_P\left[\exp\left((u+iw)^{\text{tr}}(L_t - L_s) - (t-s)\Psi(u)\right)\middle|\ \mathcal{F}_s\right] \\
&= E_P\left[Z_{t-s}^u \exp\left(i\ w^{\text{tr}}(L_{t-s})\right)\right] \\
&= E_{Q^u}\left[\exp\left(i\ w^{\text{tr}} L_{t-s}\right)\right],
\end{aligned}$$

so that L is a Q^u-Lévy process by Proposition 1.7. $\square$

Lemma 4.3 *Let $u \in \mathbb{R}^d$ with $E_P[\exp(u^{\text{tr}} L_1)] < \infty$ and let Q^u be an Esscher martingale measure for UL. If $u \in \text{Im}(U^{\text{tr}})$, then the entropy process of Q^u with respect to P is finite-valued and given by $I_t(Q^u|P) = -t\Psi(u)$.*

Proof. L is a Lévy process under Q^u, so UL is a Lévy process by Corollary 1.35. Thus UL is a local martingale and Lévy process under Q^u, and hence a true Q^u-martingale, and if we write $u = U^{\text{tr}}\tilde{u}$, we have

$$I_t(Q^u|P) = E_{Q^u}\left[\log Z_t^u\right] = E_{Q^u}\left[\tilde{u}^{\text{tr}} U L_t - t\Psi(u)\right] = -t\Psi(u).$$

However $\Psi(u)$ is well-defined (i.e. in $\mathbb{R}$) by Theorem A.7. Note that obviously $\Psi(u) \leq 0$ since $I_t(Q^u|P) \geq 0$. $\square$

The following criterion is a Lévy version of Proposition 3.2 in Grandits and Rheinländer (2002), formulated for an infinite time horizon. Thereby we see that the Esscher martingale measure for UL is optimal in $\mathcal{Q}_\ell^U$, provided it exists. Note that here we do not assume that Q^E exists.

Proposition 4.4 *Let L be a Lévy process. If there exists an Esscher martingale measure Q^u for UL with $u \in \operatorname{Im}\left(U^{\mathrm{tr}}\right)$, i.e. a measure $Q^u \in \mathcal{Q}_e^U \cap \mathcal{Q}_f^U \cap \mathcal{Q}_\ell^U$ with density process*

$$Z_t^u = \left.\frac{dQ^u}{dP}\right|_{\mathcal{F}_t} = \frac{\exp\left(u^{\mathrm{tr}} L_t\right)}{E_P\left[\exp\left(u^{\mathrm{tr}} L_t\right)\right]}$$

for some $u \in \operatorname{Im}\left(U^{\mathrm{tr}}\right)$, then $I_t(Q^u|P) \leq I_t(R|P)$ for all $R \in \mathcal{Q}_\ell^U$ and therefore $Q^u = Q_\ell^E(UL)$.

Proof. Let $u = U^{\mathrm{tr}}\tilde{u}$. From Lemma 4.3 we know that

$$I_t(Q^u|P) = -t\Psi(u),$$

and for an arbitrary measure $R \in \mathcal{Q}_\ell$ we have

$$\begin{aligned}
I_t(R|P) &= I_t(R|Q^u) + E_R\left[\log Z_t^u\right] \\
&= I_t(R|Q^u) + E_R\left[u^{\mathrm{tr}} L_t\right] - t\Psi(u) \\
&\geq E_R\left[\tilde{u}^{\mathrm{tr}} U L_t\right] - t\Psi(u) \\
&= -t\Psi(u) \\
&= I_t(Q^u|P),
\end{aligned}$$

hence the claim. □

In order to decide whether a given Esscher measure Q^u is a martingale measure for L we use Proposition 3.7. To that end we need to know the Girsanov quantities of Q^u.

Proposition 4.5 *Let $u \in \mathbb{R}^d$ with $E_P[\exp(u^{\mathrm{tr}} L_1)] < \infty$ and let Q^u be an Esscher measure. Then the Girsanov quantities of Q^u are given by $\beta^u = u$, $Y^u(x) = e^{u^{\mathrm{tr}} x}$.*

Proof. Let β^u and Y^u be the Girsanov quantities associated with Q^u by Theorem 2.2. By Proposition 4.2 L is a Lévy process under Q^u, so Theorem 1.27 yields that β^u is a constant and $Y^u = Y^u(x)$ is a deterministic function. Corollary 2.9 then yields $Z_t^u = \mathcal{E}(N^u)_t$ with

$$N_t^u = (\beta^u)^{\mathrm{tr}} L_t^c + (Y^u - 1) * (\mu^L - \nu^P)_t.$$

So by the explicit formula for the stochastic exponential we get

$$\log Z_t^u = (\beta^u)^{\mathrm{tr}} L_t^c - \frac{1}{2}(\beta^u)^{\mathrm{tr}} c \beta^u\ t + (Y^u - 1) * (\mu^L - \nu^P)_t + \sum_{s \leq t}\left(\log(1 + \Delta N_s^u) - \Delta N_s^u\right).$$

On the other hand, by Theorem A.7 we know that

$$Z_t^u = \exp\left(u^{\mathrm{tr}} L_t - t\ \Psi(u)\right). \tag{4.1}$$

Now under P we have the decomposition $L = L^c + V$, where V is the sum of a purely discontinuous local P-martingale and a process of finite variation (this follows immediately from Jacod and Shiryaev (1987), Proposition I.4.27). Then (4.1) yields

$$\log Z_t^u = u^{\mathrm{tr}} L_t^c + u^{\mathrm{tr}} V_t - t\ \Psi(u),$$

so $\log Z^u$ is a semimartingale with continuous P-martingale part $u^{\mathrm{tr}} L^c$. Comparing the two decompositions of $\log Z^u_t$, we see that $\beta^u = u$.

To obtain the explicit form of Y^u we examine the jumps of Z^u. On the one hand, from $dZ^u = Z^u_- \, dN^u$, these are given by

$$\Delta Z^u_t = Z^u_{t-} \Delta N^u_t = Z^u_{t-}(Y^u(\Delta L_t) - 1)$$

(note that $Y^u(0) = 1$ by Remark 2.4), and on the other hand, we have

$$\Delta Z^u_t = \left(e^{u^{\mathrm{tr}} L_t} - e^{u^{\mathrm{tr}} L_{t-}} \right) \; e^{-t \; \Psi(u)} = Z^u_{t-} \left(e^{u^{\mathrm{tr}} \Delta L_t} - 1 \right).$$

So since Z^u_- is strictly positive, we see that $Y^u(x) = e^{u^{\mathrm{tr}} x}$. $\square$

4.3 A Generalization of the Black-Scholes Model

Recall that for the results about the preservation of the Lévy property under the entropy-minimizing martingale measure Q^E in Chapter 3 we have always assumed that Q^E exists. Here we state some sufficient conditions on the P-Lévy characteristics of L which ensure the existence of Q^E.

Let $L = (L^i)_{1 \le i \le d}$ be a d-dimensional P-Lévy process with respect to the P-augmentation of the filtration it generates, and assume that $\Delta L^i > -1$, i.e. the jumps of every component of L are strictly bigger than -1. Let (b, c, K) be the P-Lévy characteristics of L relative to some fixed truncation function h.

Let $X^i = \mathcal{E}(L^i)$, $i \in \{1, \dots, d\}$ be geometric Lévy processes, i.e. $dX^i = X^i_- \, dL^i$, or

$$X^i_t = \exp\left(L^i_t - \frac{1}{2} \langle L^{i,c} \rangle_t \right) \prod_{s \le t} (1 + \Delta L^i_s) e^{-\Delta L^i_s}.$$

Note that if $d = 1$ and if we set $L_t = \sigma W_t + \mu t$ for a P-Brownian motion W, $\sigma > 0$ and $\mu \in \mathbb{R}$, then X is geometric Brownian motion, so our model is indeed a generalization of the Black-Scholes model. Since we consider multidimensional processes with an infinite time horizon, our model is also a generalization of the model of Fujiwara and Miyahara (2003). They consider the setting of a one-dimensional *exponential Lévy process*, i.e. $X = e^{\tilde{L}}$. Modeling stocks by exponential Lévy processes has the advantage that one does not need to require the jumps of the Lévy process to be bounded from below. However our model allows us to use the fact stated in the next proposition, namely that X is a local martingale if and only if L is a local martingale, which obviously simplifies many arguments. In Appendix D we give a detailed comparison of our results with those of Fujiwara and Miyahara (2003) and also point out some issues which are not presented in full perspicuousness there.

We are looking for the entropy-minimizing martingale measure for X, and the next proposition shows that with the above conditions on L we can confine ourselves to the study of the entropy-minimizing martingale measure for L instead of X. Note that in the above case U is the identity matrix, so we omit the dependence on U in the sets of martingale measures.

Proposition 4.6 *In view of Definition 1.14 we have for the above X*

$$\mathcal{Q}_x(X) = \mathcal{Q}_x(L) =: \mathcal{Q}_x \qquad \textit{for } x \in \{a, e, f\}.$$

For simplicity we also write $\mathcal{Q}_\ell := \mathcal{Q}_\ell(L)$.

Proof. If L is a local martingale, then the stochastic exponential of L^i is also a local martingale (cf. Theorem 2.1). On the other hand, if $X^i = \mathcal{E}(L^i)$, then L^i may be written as $\int \frac{1}{X^i_-}\, dX^i$ and the integrand is locally bounded since X^i_- is left-continuous and strictly positive. Thus if the integrator is a local martingale, so is the integral. □

From Proposition 4.6 it is immediate that $Q^E(X) = Q^E(L) =: Q^E$. Furthermore the informal argument in Section 4.1 yields that Q^E should be an Esscher martingale measure for L, and we get the following characterization of the entropy-minimizing martingale measure for X.

Theorem 4.7 *Suppose there exists $u^* \in \mathbb{R}^d$ such that*

$$\begin{aligned}
&\textit{(i)} \qquad \int_{\mathbb{R}^d} \left| x e^{(u^*)^{\mathrm{tr}} x} - h(x) \right| \ K(dx) < \infty, \\
&\textit{(ii)} \qquad b + c u^* + \int_{\mathbb{R}^d} \left(x e^{(u^*)^{\mathrm{tr}} x} - h(x) \right) \ K(dx) = 0.
\end{aligned}$$

Then there exists $Q^* \in \mathcal{Q}_e \cap \mathcal{Q}_f \cap \mathcal{Q}_\ell$ with density process Z^*, given by

$$Z^*_t = \frac{\exp\left((u^*)^{\mathrm{tr}} L_t\right)}{E_P\left[\exp\left((u^*)^{\mathrm{tr}} L_t\right)\right]}, \qquad t \geq 0.$$

Moreover, Q^E exists and $Q^* = Q^E$.

Proof. We have $E_P\left[\exp\left((u^*)^{\mathrm{tr}} L_t\right)\right] < \infty$ if and only if $\int_{\{|x|>1\}} e^{(u^*)^{\mathrm{tr}} x} \ K(dx) < \infty$ by Theorem A.7. But

$$\begin{aligned}
\int_{\{|x|>1\}} e^{(u^*)^{\mathrm{tr}} x} \ K(dx) &\leq \int_{\{|x|>1\}} |x| e^{(u^*)^{\mathrm{tr}} x} \ K(dx) + \int_{\{|x|\leq 1\}} \left| x \left(e^{(u^*)^{\mathrm{tr}} x} - 1 \right) \right| \ K(dx) \\
&= \int_{\mathbb{R}^d} \left| x e^{(u^*)^{\mathrm{tr}} x} - h_0(x) \right| \ K(dx) \\
&\leq \int_{\mathbb{R}^d} \left| x e^{(u^*)^{\mathrm{tr}} x} - h(x) \right| \ K(dx) + \int_{\mathbb{R}^d} |h_0(x) - h(x)| \ K(dx),
\end{aligned}$$

which is finite by condition (ii) and Lemma C.1, respectively. So Z^* is well-defined, and it is straightforward to see that Z^* is a strictly positive P-martingale. Thus by Theorem 2.10 there exists a measure Q^* with density process Z^*.

Now Q^* is an Esscher measure by Definition 4.1 and thus L is a Lévy process under Q^* by Proposition 4.2. Then by Proposition 4.5 the Girsanov quantities of Q^* are $\beta^* = u^*$ and $Y^*(x) = e^{(u^*)^{\mathrm{tr}}x}$, and conditions (i) and (ii) are the conditions from Proposition 3.7 for L to be a local Q^*-martingale, so that Q^* is an Esscher martingale measure for L, and Proposition 4.4 yields that $Q^* = Q^E_\ell$.

Finally by Lemma 4.3 we know that $I_t(Q^*|P) < \infty$ for all $t \geq 0$, and since Z^* is strictly positive, we have $Q^* \in \mathcal{Q}_e \cap \mathcal{Q}_f \cap \mathcal{Q}_\ell$, so that $\mathcal{Q}_e \cap \mathcal{Q}_f \cap \mathcal{Q}_\ell \neq \emptyset$. Then $Q^* = Q^E$ by Corollary 3.20. □

Remark 4.8 As pointed out in Section 4.1, the problem of finding the entropy-minimizing martingale measure for X (or, equivalently, L) reduces to the deterministic problem of minimizing the functional

$$\hat{I}(\beta, Y) := \frac{1}{2}\beta^{\mathrm{tr}}c\beta + \int_{\mathbb{R}^d} f(Y(x))\ K(dx)$$

over all deterministic $\beta \in \mathbb{R}^d$ and measurable $Y : \mathbb{R}^d \to \mathbb{R}_+$ subject to the constraints

$$\begin{aligned} &(i) \qquad \int_{\mathbb{R}^d} f(Y(x))\ K(dx) < \infty \\ &(ii) \qquad \int_{\mathbb{R}^d} |x\ Y(x) - h(x)|\ K(dx) < \infty \\ &(iii) \qquad b + c\beta + \int_{\mathbb{R}^d} (x\ Y(x) - h(x))\ \ K(dx) = 0. \end{aligned}$$

Obviously (i) is the condition for finite relative entropy (and also finiteness of the functional $\hat{I}$), whereas (ii) and (iii) come from the martingale condition. Now let $\mathcal{H}$ be the set of all (β, Y) which fulfill (i), (ii) and (iii) and define $Y^u(x) = e^{u^{\mathrm{tr}}x}$. From the informal argument in Section 4.1 we have the candidate tuple (u, Y^u), and we show

$$\hat{I}(u, Y^u) \leq \hat{I}(\beta, Y)$$

for all $(\beta, Y) \in \mathcal{H}$.

With these arguments in mind, the following theorem now gives an analytic proof of the optimality of the Esscher martingale measure for L in the set of all Lévy martingale measures. ◇

Theorem 4.9 *Suppose there exists $u \in \mathbb{R}^d$ such that $(u, Y^u) \in \mathcal{H}$. Then $\hat{I}(u, Y^u) \leq \hat{I}(\beta, Y)$ for all $(\beta, Y) \in \mathcal{H}$.*

Proof. We claim that

$$\begin{aligned} 0 \ &\leq\ \frac{1}{2}(\beta - u)^{\mathrm{tr}}c(\beta - u) + \int_{\mathbb{R}^d} Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)\ K(dx) \\ &=\ \frac{1}{2}\beta^{\mathrm{tr}}c\beta + \int_{\mathbb{R}^d} f(Y(x))\ K(dx) - \left(\frac{1}{2}u^{\mathrm{tr}}cu + \int_{\mathbb{R}^d} f\left(Y^u(x)\right)\ K(dx)\right). \end{aligned} \tag{4.2}$$

Nonnegativity of $\frac{1}{2}(\beta - u)^{\mathrm{tr}}c(\beta - u) + \int Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)\ K(dx)$ is immediate since c is nonnegative definite and $Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right) \geq 0$. Concerning K-integrability of $Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)$ we see by a simple calculation that for $x \in \mathbb{R}^d$

$$Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right) = f(Y(x)) - f(Y^u(x)) + u^{\mathrm{tr}}x(Y^u(x) - Y(x)), \tag{4.3}$$

where $u^{\mathrm{tr}}x(Y^u(x) - Y(x))$ is K-integrable by condition (ii) for Y^u and Y, since

$$\left|u^{\mathrm{tr}}x(Y^u(x) - Y(x))\right| \leq \left|u^{\mathrm{tr}}xY^u(x) - h(x))\right| + \left|u^{\mathrm{tr}}xY(x) - h(x)\right|.$$

Let us now show the equality in (4.2). We have

$$(\beta - u)^{\mathrm{tr}}c(\beta - u) = \beta^{\mathrm{tr}}c\beta + u^{\mathrm{tr}}cu - 2u^{\mathrm{tr}}c\beta,$$

where

$$u^{\mathrm{tr}}c\beta = -\left(u^{\mathrm{tr}}b + u^{\mathrm{tr}}\int_{\mathbb{R}^d}(xY(x) - h(x))\ K(dx)\right)$$

by condition (iii) for (β, Y), and

$$u^{\mathrm{tr}}b = -\left(u^{\mathrm{tr}}cu + u^{\mathrm{tr}}\int_{\mathbb{R}^d}(xY^u(x) - h(x))\ K(dx)\right)$$

by condition (iii) for (u, Y^u). This yields

$$u^{\mathrm{tr}}c\beta = u^{\mathrm{tr}}cu - \int_{\mathbb{R}^d} u^{\mathrm{tr}}x\left(Y(x) - Y^u(x)\right)\ K(dx),$$

and thus

$$(\beta - u)^{\mathrm{tr}}c(\beta - u) = \beta^{\mathrm{tr}}c\beta - u^{\mathrm{tr}}cu + 2\int_{\mathbb{R}^d} u^{\mathrm{tr}}x\left(Y(x) - Y^u(x)\right)\ K(dx). \tag{4.4}$$

Furthermore by (4.3) we have

$$\begin{aligned}
&\int_{\mathbb{R}^d} Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)\ K(dx)\\
(4.5)\quad = &\int_{\mathbb{R}^d} f(Y(x))\ K(dx) - \int_{\mathbb{R}^d} f(Y^u(x))\ K(dx) + \int_{\mathbb{R}^d} u^{\mathrm{tr}}x(Y^u(x) - Y(x))\ K(dx).
\end{aligned}$$

So putting (4.4) and (4.5) together we obtain

$$\begin{aligned}
&\frac{1}{2}(\beta - u)^{\mathrm{tr}}c(\beta - u) + \int_{\mathbb{R}^d} Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)\ K(dx)\\
= &\frac{1}{2}\beta^{\mathrm{tr}}c\beta + \int_{\mathbb{R}^d} f(Y(x))\ K(dx) - \left(\frac{1}{2}u^{\mathrm{tr}}cu + \int_{\mathbb{R}^d} f\left(Y^u(x)\right)\ K(dx)\right),
\end{aligned}$$

which yields the claim. □

Remark 4.10 The proof of Theorem 4.9 is closely related to the proof of Proposition 4.4. There we show that

$$I_t(R|P) = I_t(R|Q) + I_t(Q|P),$$

where Q is the Esscher martingale measure for L and R is an arbitrary Lévy martingale measure. In the proof of Theorem 4.9 this equality corresponds to (4.2). We have $I_t(R|P) = \hat{I}(\beta, Y)\ t$ and $I_t(Q|P) = \hat{I}(u, Y^u)\ t$. Heuristically it is easy to see that the Girsanov quantities of R with respect to Q relative to L are given by $\tilde{\beta} = \beta - u$ and $\tilde{Y} = \frac{Y}{Y^u}$, so that the entropy process of R with respect to Q is actually given by

$$I_t(R|Q) = \left(\frac{1}{2} \tilde{\beta}^{\mathrm{tr}} c \tilde{\beta} + \int_{\mathbb{R}^d} f(\tilde{Y}(x))\ \tilde{K}(dx) \right) t,$$

where $\tilde{K}(dx) = Y^u(x)\ K(dx)$ is the third Lévy characteristic of L with respect to Q. $\diamond$

4.4 Stochastic Volatility Models

We finally come to the point where the matrix U comes into play. We deal with the following problem (after a suggestion by D. Becherer): Let $L = (L^1, L^2)$ be a 2-dimensional Lévy process under P with P-Lévy characteristics (b, c, K) and let X be the solution of

$$dX_t = \sigma(X_{t-}, L^2_{t-}, t)\ dL^1_t$$

for some measurable $\sigma\colon \mathbb{R}^2 \times \mathbb{R}_+ \to \mathbb{R}$ such that a strictly positive solution X exists and such that X is a local martingale if and only if L^1 is a local martingale. Basically, this reduces to the problem of optimal change of measure, where we only require L^1 to be a local martingale. More precisely, we require $U^1 L$ to be a local martingale, where

$$U^1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

is the projection matrix on the first coordinate. Now, if L^1 and L^2 are independent, the optimal change of measure should leave L^2 unchanged, but the more interesting question is what happens if L^1 and L^2 are correlated.

Note that in general the correlation of L^1 and L^2 is not only manifest in c, which is the covariance matrix of the Brownian parts of L, but also in the "jump measure" K. In fact, L^1 and L^2 are independent under P if and only if c is a diagonal matrix and K is concentrated on the coordinate axes, cf. Bertoin (1996), Exercise 5.1.

If we have the existence of $Q^E := Q^E(U^1 L)$, i.e. the entropy-minimizing martingale measure for $U^1 L$, we know from Theorem 3.19 that L is a Lévy process under Q^E, and Corollary 3.20 shows that $Q^E \in \mathcal{Q}_e^{U^1} \cap \mathcal{Q}_f^{U^1} \cap \mathcal{Q}_l^{U^1}$, so the problem of finding Q^E reduces to the following deterministic problem.

Define for $\beta \in \mathbb{R}^2$, $Y\colon \mathbb{R}^2 \to \mathbb{R}_+$ the functional

$$\hat{I}(\beta, Y) = \frac{1}{2}\beta^{\mathrm{tr}} c\beta + \int_{\mathbb{R}^2} f(Y(x))\ K(dx)$$

and let $\mathcal{H}^1$ be the set of all (β, Y) with a constant $\beta \in \mathbb{R}^2$ and a measurable function $Y\colon \mathbb{R}^2 \to \mathbb{R}_+$ which satisfy

$$\begin{array}{ll} (i) & \displaystyle\int_{\mathbb{R}^2} f(Y(x))\ K(dx) < \infty \\ (ii) & \displaystyle\int_{\mathbb{R}^2} |x^1\ Y(x) - h^1(x)|\ K(dx) < \infty \\ (iii) & \displaystyle b^1 + c^{11}\beta^1 + c^{12}\beta^2 + \int_{\mathbb{R}^2} \left(x^1 Y(x) - h^1(x)\right)\ K(dx) = 0. \end{array}$$

Again these conditions come from the finiteness of the relative entropy and the martingale condition for $U^1 L$.

Proposition 4.4 suggests that the optimal Lévy martingale measure for U^1L is an Esscher martingale measure for U^1L with $u \in \mathrm{Im}((U^1)^{\mathrm{tr}})$, i.e. $u^2 = 0$. In the sequel, however, we show how to obtain the optimal Girsanov quantities analytically.

Assume $c^{11} \neq 0$, then by condition (iii)

$$\beta^1 = -\frac{1}{c^{11}}\left(b^1 + c^{12}\beta^2 + k(Y)\right), \tag{4.6}$$

where $k(Y) = \int_{\mathbb{R}^2}(x^1Y(x) - h^1(x))\ K(dx)$. So for fixed $\beta^2 \in \mathbb{R}$, minimizing $\hat{I}$ under the constraint (iii) reduces to minimizing

$$\begin{aligned} \hat{I}(\beta, Y) \quad &=: \quad \bar{I}(Y) \\ &= \quad \frac{1}{2c^{11}}\left((b^1)^2 + \det c\ (\beta^2)^2 + 2b^1k(Y) + (k(Y))^2\right) + \int_{\mathbb{R}^2} f(Y(x))\ K(dx), \end{aligned} \tag{4.7}$$

and with a similar variational argument as in Section 4.1 we get the "necessary condition"

$$Y(x) = \exp\left(\frac{c^{11}\beta^1 + c^{12}\beta^2}{c^{11}} x^1\right)$$

for the Girsanov quantities of the optimal measure for fixed β^2. It then remains to find the optimal β^2. However (4.7) suggests $\beta^2 = 0$, and this together with (4.6) yields the optimal tuple (β, Y), as the following theorem shows. As in Section 4.3 we define $Y^u(x) = e^{u^{\mathrm{tr}}x}$ for $u \in \mathbb{R}^2$.

Theorem 4.11 *Suppose there exists $u \in \mathbb{R}^2$ with $u^2 = 0$ and such that $(u, Y^u) \in \mathcal{H}^1$. Then $\hat{I}(u, Y^u) \leq \hat{I}(\beta, Y)$ for all $(\beta, Y) \in \mathcal{H}^1$.*

Proof. Similarly to the proof of Theorem 4.9 we show that

$$\begin{aligned} 0 &\leq \frac{1}{2}(\beta - u)^{\text{tr}}c(\beta - u) + \int_{\mathbb{R}^2} Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)\ K(dx) \\ &= \frac{1}{2}\beta^{\text{tr}}c\beta + \int_{\mathbb{R}^2} f(Y(x))\ K(dx) - \left(\frac{1}{2}u^{\text{tr}}cu + \int_{\mathbb{R}^2} f\left(Y^u(x)\right)\ K(dx)\right), \end{aligned}$$

where integrability of $Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)$ with respect to K is shown as in the proof of Theorem 4.9.

Using condition (iii) for (β, Y) and (u, Y^u), and with the fact that $u^2 = 0$, a simple calculation shows

$$\frac{1}{2}(\beta - u)^{\text{tr}}c(\beta - u) = \frac{1}{2}\beta^{\text{tr}}c\beta - \frac{1}{2}u^{\text{tr}}cu - u^1 \int_{\mathbb{R}^2} x^1(Y^u(x) - Y(x))\ K(dx).$$

Integrability of $x^1(Y^u(x) - Y(x))$ follows from

$$|x^1(Y^u(x) - Y(x))| \leq |x^1 Y^u(x) - h^1(x)| + |x^1 Y(x) - h^1(x)|$$

and condition (ii). On the other hand

$$\begin{aligned} Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right) &= f(Y(x)) - u^1 x^1 Y(x) + Y^u(x) - 1 \\ &= f(Y(x)) - f(Y^u(x)) + u^1 x^1(Y^u(x) - Y(x)), \end{aligned}$$

so altogether we have

$$\begin{aligned} &\frac{1}{2}(\beta - u)^{\text{tr}}c(\beta - u) + \int_{\mathbb{R}^d} Y^u(x)\ f\left(\frac{Y(x)}{Y^u(x)}\right)\ K(dx) = \\ &= \frac{1}{2}\beta^{\text{tr}}c\beta - \frac{1}{2}u^{\text{tr}}cu + \int_{\mathbb{R}^2} f(Y(x))\ K(dx) - \int_{\mathbb{R}^2} f(Y^u(x))\ K(dx), \end{aligned}$$

the claimed result. □

Corollary 4.12 *Suppose there exists $u \in \mathbb{R}^2$ with $u^2 = 0$ and such that $(u, Y^u) \in \mathcal{H}^1$. Then the entropy-minimizing martingale measure for U^1L exists and is given by*

$$\left.\frac{dQ^E(U^1L)}{dP}\right|_{\mathcal{F}_t} = \frac{\exp\left(u^1 L_t^1\right)}{E_P[\exp\left(u^1 L_t^1\right)]}.$$

Proof. As in the proof of Theorem 4.9 $(u, Y^u) \in \mathcal{H}^1$ implies $E_P[\exp\left(u^1 L_t^1\right)] < \infty$, so $Q^E(U^1L)$ as defined above exists and is the entropy-minimizing Lévy martingale measure by Theorem 4.11. Furthermore $Q \in \mathcal{Q}_e^{U^1} \cap \mathcal{Q}_f^{U^1} \cap \mathcal{Q}_l^{U^1}$, so that $\mathcal{Q}_e^{U^1} \cap \mathcal{Q}_f^{U^1} \cap \mathcal{Q}_l^{U^1} \neq \emptyset$, and thus Q is the entropy-minimizing martingale measure for UL by Corollary 3.20. □

Remark 4.13 Theorem 4.11 and Corollary 4.12 show that the entropy-minimizing measure among all measures under which U^1L is a local martingale coincides with the Esscher martingale measure for U^1L. Loosely speaking we take the entropy-minimizing martingale measure for L^1 separately, whereas the change in L^2 only results from the correlation of L^1 and L^2. However note that by Theorem 1.34 the characteristics of L^1 and L^2 are not simply the projections of the characteristics of L. $\diamond$

Part III

CONVERGENCE RESULTS IN AN INCOMPLETE MODEL

Chapter 5

Model Convergence

This chapter is concerned with the approximation of a financial market model $(S, \eta, \mathbb{F})$, where $S = (S^i)_{i \in \{1,\dots,d\}}$ is a tradable asset, η is a process with values in a discrete set, by which we model a non-tradable factor of risk and $\mathbb{F}$ is a filtration generated by an r-dimensional Brownian motion and a multivariate point process, which drive S and η, respectively. In addition we assume mutual dependences of S and η. The impact of η on the evaluation of S is manifest in the assumption that the current value of η influences the dynamics of S, whereas the values of S influence the dynamics of η. For instance we can think of η as the rating given by some agency to an asset on a stock market, whose evolution is itself influenced by the rating results. The main problem one faces in this setting is to model these mutual dependences. The model is taken from Section 3 of Becherer and Schweizer (2003), who also give a recipe to construct such a model by means of a suitable change of measure. Notice that if we suppose $r > d$ the incompleteness of the model results not only from the fact that there are more driving Brownian motions than tradable assets, a feature which is well-understood in mathematical finance, but is crucially due to the non-tradable factors of risk.

In Section 5.1 we present the model in continuous time and also give some details on how to obtain existence and uniqueness of (S, η). In Section 5.2 we construct an analogous discrete-time model (S^n, η^n), and we show that $(S^n, \eta^n) \xrightarrow{\mathcal{L}} (S, \eta)$ if η is an autonomous process (i.e. its evolution is not influenced by S). Section 5.3 then shows that convergence still holds after a change of measure, and finally we give in Section 5.4 a detailed analysis of the dynamics of η^n and their influence on the change of measure between P^n and P'^n. This will be important for the next chapter where we give an approximation result for a backward stochastic differential equation driven by S and η, which will yield convergence for endogenous price processes of path-dependent contingent claims that depend in addition on the changes of the non-tradable factors of risk.

5.1 A Model with Untradable Risk

In this section we adopt the basic model from Becherer and Schweizer (2003), Section 3. Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a stochastic basis and let (S, η) be the solution of the following system of stochastic differential equations with values in $\mathbb{R}^d \times \{1, \dots, m\}$, where $d, r, m \in \mathbb{N}$:

$$dS_t = \Gamma(t, S_t, \eta_{t-})\, dt + \Sigma(t, S_t, \eta_{t-})\, dW_t, \quad S_0 \in \mathbb{R}^d, \tag{5.1}$$

$$d\eta_t = \sum_{\ell,j=1}^{m} (j - \ell) \mathbb{1}_{\{\eta_{t-}=\ell\}}\, dN_t^{\ell j}, \quad \eta_0 \in \{1, \dots, m\}, \tag{5.2}$$

where Γ and Σ are $\mathbb{R}^d$-valued respectively $\mathbb{R}^{d\times r}$-valued functions which are C^1 with respect to (t, x). Furthermore $W = (W^i)_{i\in\{1,\dots,r\}}$ is an $\mathbb{R}^r$-valued $(P, \mathbb{F})$-standard Brownian motion and $N = (N^{\ell j})_{\ell,j\in\{1,\dots,m\}}$ is a multivariate $\mathbb{F}$-adapted point process such that

$$N_t^{\ell j} \text{ has } (P, \mathbb{F})\text{-intensity } \lambda^{\ell j}(t, S_t) \text{ for } \ell, j \in \{1, \dots, m\}, \tag{5.3}$$

where $\lambda^{\ell j} \colon [0, T] \times \mathbb{R}^d \to [0, \infty)$ are bounded functions of class C^1 with bounded gradients.

Note that η jumps from ℓ to j whenever $N^{\ell j}$ jumps by 1; hence if $\eta_0 \in \{1, \dots, m\}$, then η never leaves $\{1, \dots, m\}$. So it would be sufficient to assume that $\Gamma(t, x, \cdot)$ and $\Sigma(t, x, \cdot)$ are defined only on the set $\{1, \dots, m\}$. However for the approximation of η by a suitable discrete-time process it turns out that the approximating processes η^n may leave $\{1, \dots, m\}$. So it will be convenient to introduce an additional cemetery point π in the range of η_t and to assume that $\Gamma(t, x, \cdot)$ and $\Sigma(t, x, \cdot)$ are defined on the set $I_m := \{\pi\} \cup \{1, \dots, m\}$.

The idea to construct such a model with mutual dependences between the processes involved is to first consider a solution (S, η) to (5.1) and (5.2) in the case where $N = (N^{\ell j})$ is a standard multivariate Poisson process independent of W, and then obtain the desired dependences by a suitable change of measure.

Since the overall idea of approximating (S, η) by discrete-time processes is similar to the induction argument for existence of a unique solution of (5.1) in Example 3.3.3 in Becherer (2001), we give the following details. Let W be an r-dimensional standard Brownian motion and $N = (N^{\ell j})$ a standard multivariate Poisson process independent of W on some stochastic basis $(\Omega, \mathcal{F}', \mathbb{F}', P')$, where $\mathcal{F}'$ is P'-complete and $\mathbb{F}'$ satisfies the usual conditions. In order to get a unique strong solution to the system (5.1), (5.2) we assume that for each $\ell \in I_m$ the functions $x \mapsto \Gamma(t, x, \ell)$ and $x \mapsto \Sigma(t, x, \ell)$ are globally Lipschitz continuous in x, uniformly in $t \in [0, T]$. Furthermore we define the jump times of η by

$$\begin{cases} \tau_0 = 0, \\ \tau_k = \inf\left\{t > \tau_{k-1} \mid |\Delta\eta_t| > \frac{1}{2}\right\} \wedge T. \end{cases} \tag{5.4}$$

Now the existence of a unique solution S of (5.1) is shown by the following induction argument. Let $S_t^{(0)} \equiv S_0$ and suppose that $S^{(k)}$ is the unique solution of (5.1) on $[\![0, \tau_k]\!]$ for some $k \geq 0$.

Then by Protter (1990), Theorem V.3.7, the stochastic differential equation

$$(5.5)\quad S_t^{(k+1)} = S_{t\wedge\tau_k}^{(k)} + \int_0^t \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}\ \Gamma(s, S_{s-}^{(k+1)}, \eta_{\tau_k})\ ds + \int_0^t \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}\ \Sigma(s, S_{s-}^{(k+1)}, \eta_{\tau_k})\ dW_s$$

has a unique solution $S^{(k+1)}$, and $S^{(k+1)}$ solves (5.1) on $[\![0,\tau_{k+1}]\!]$. Since $P[\tau_k \geq T] \to 1$ as $n \to \infty$,

$$S_t = S_0 + \sum_{k=1}^{\infty} S_t^{(k)}\ \mathbb{1}_{]\!]\tau_{k-1},\tau_k]\!]}, \quad t \in [0,T]$$

is the unique solution of (5.1).

In order to show convergence of an analogously defined discrete-time model we further dissect S in order to get deterministic functions as coefficients in (5.5) so that we can apply results from Kurtz and Protter (1996) concerning convergence of solutions of stochastic differential equations. To that end recall that π is an additional cemetery point, and let $S^{(k)}$ be given as above. Let $S_t^{(0,\ell)} \equiv S_0$ for $\ell \in I_m$ and define recursively for all $k \in \{0,1,\dots\}$ and $\ell \in I_m$ the process $S^{(k+1,\ell)}$ as the unique solution of

$$(5.6)\quad S_t^{(k+1,\ell)} = S_{t\wedge\tau_k}^{(k)} + \int_0^t \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}\ \Gamma(s, S_{s-}^{(k+1,\ell)}, \ell)\ ds + \int_0^t \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}\ \Sigma(s, S_{s-}^{(k+1,\ell)}, \ell)\ dW_s.$$

Furthermore define for $\ell \in I_m$ the process $S^{(k),\ell}$ by

$$S_t^{(k),\ell} = \int_0^t \Gamma(s, S_{s-}^{(k)}, \ell)\ ds + \int_0^t \Sigma(s, S_{s-}^{(k)}, \ell)\ dW_s, \quad t \in [0,T],$$

Lemma 5.1 *Let the processes $S^{(k)}$, $S^{(k,\ell)}$ and $S^{(k),\ell}$ be given as above. Then the stochastic differential equation (5.6) may be written as*

$$(5.7)\quad S_t^{(k+1,\ell)} = S_{t\wedge\tau_k}^{(k)} - S_{t\wedge\tau_k}^{(k),\ell} + \int_0^{t\wedge\tau_{k+1}} \Gamma(s, S_{s-}^{(k+1,\ell)}, \ell)\ ds + \int_0^{t\wedge\tau_{k+1}} \Sigma(s, S_{s-}^{(k+1,\ell)}, \ell)\ dW_s.$$

Proof. By the definition of $S^{(k,\ell)}$ in (5.6) we have

$$S_t^{(k+1,\ell)} - S_{t\wedge\tau_k}^{(k)} = \int_0^t \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}(s)\ \Gamma(s, S_{s-}^{(k+1,\ell)}, \ell)\ ds + \int_0^t \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}(s)\ \Sigma(s, S_{s-}^{(k+1,\ell)}, \ell)\ dW_s,$$

or

$$d(S_t^{(k+1,\ell)} - S_{t\wedge\tau_k}^{(k)}) = \mathbb{1}_{]\!]\tau_k,\tau_{k+1}]\!]}(t) \left(\Gamma(t, S_{t-}^{(k+1,\ell)}, \ell)\ dt + \Sigma(t, S_{t-}^{(k+1,\ell)}, \ell)\ dW_t\right).$$

This yields

$$\begin{aligned}
& d\left(S_t^{(k+1,\ell)} - S_{t\wedge\tau_k}^{(k)} - \int_0^{t\wedge\tau_{k+1}} \Gamma(s, S_{s-}^{(k+1,\ell)}, \ell)\ ds - \int_0^{t\wedge\tau_{k+1}} \Sigma(s, S_{s-}^{(k+1,\ell)}, \ell)\ dW_s\right) \\
= \ & -\mathbb{1}_{]\!]0,\tau_k]\!]}(t) \left(\Gamma(t, S_{t-}^{(k+1,\ell)}, \ell)\ dt + \Sigma(t, S_{t-}^{(k+1,\ell)}, \ell)\ dW_t\right) \\
= \ & -\mathbb{1}_{]\!]0,\tau_k]\!]}(t) \left(\Gamma(t, S_{t-}^{(k)}, \ell)\ dt + \Sigma(t, S_{t-}^{(k)}, \ell)\ dW_t\right) \\
= \ & -dS_t^{(k),\ell}\mathbb{1}_{]\!]0,\tau_k]\!]}(t),
\end{aligned}$$

since $S^{(k+1,\ell)} = S^{(k)}$ on $]\!]0,\tau_k]\!]$ by construction. $\square$

Now by construction of $S^{(k+1,\ell)}$ we have

$$S^{\tau_{k+1}} = S^{(k+1)} = \sum_{\ell \in I_m} S^{(k+1,\ell)} \; \mathbb{1}_{\{\eta_{\tau_k}=\ell\}}.$$

In fact by construction $S^{(k+1,\ell)}$ coincides with $S^{(k+1)}$ on $[\![0,\tau_k]\!]$, and on $\{\eta_{\tau_k} = \ell\} \cap]\!]\tau_k,\tau_{k+1}]\!]$ the dynamics of $S^{(k+1)}$ are the same as those of $S^{(k+1,\ell)}$.

In the last step one constructs N with (5.3) by a suitable change of measure. More precisely, for $P \ll P'$ defined by

$$dP := \mathcal{E}\left(\sum_{\ell,j=1}^{m} \int \left(\lambda^{\ell j}(s,S_s) - 1\right)\left(dN_s^{\ell j} - ds\right)\right)_T dP' \tag{5.8}$$

and for the P-completion $(\Omega,\mathcal{F},\mathbb{F},P)$ of $(\Omega,\mathcal{F}',\mathbb{F}',P)$, W is still a standard Brownian motion and N has the desired properties. For details see Becherer (2001), Section 3.2 and Brémaud (1981), Theorems VI.2.3 and VI.2.4. Note that the process

$$Z := \mathcal{E}\left(\sum_{\ell,j=1}^{m} \int \left(\lambda^{\ell j}(s,S_s) - 1\right)\left(dN_s^{\ell j} - ds\right)\right) \tag{5.9}$$

is the density process of P with respect to P'.

5.2 Convergence in the Case of Independent Driving Processes

The aim in this section is to approximate the model (5.1)–(5.2) under P', where W and N are independent and N is a multivariate *standard* Poisson process, by a suitable sequence of discrete-time models. To that end we define discrete-time processes (W^n, N^n) converging to (W,N) in distribution and such that the sequence $(W^n,N^n)_{n\in\mathbb{N}}$ is good in the sense of Kurtz and Protter (1996), Definition 7.3. Then we define (S^n,η^n) as the solutions of a discrete-time version of (5.1) and (5.2), and corresponding stopping times τ_k^n like in (5.4) as well as processes $S^{(k),n}$, $S^{(k,\ell),n}$ and $S^{(k),\ell,n}$ in analogy to $S^{(k)}$, $S^{(k,\ell)}$ and $S^{(k),\ell}$. The idea to show convergence is an induction argument similar to the existence argument for S above.

For $n \in \mathbb{N}$ let $(\Omega^n,\mathcal{F}^n,P'^n)$ be a probability space and let $\left\{\xi_k^{ni},\zeta_k^{n\ell j}\right\}$ for $i \in \{1,\dots,r\}$, $\ell,j \in \{1,\dots,m\}$, $k \in \{1,\dots,n\}$ be independent under P'^n with

$$\begin{aligned} P'^n\left[\xi_k^{ni} = \pm\sqrt{\frac{T}{n}}\right] &= \frac{1}{2}, \quad i \in \{1,\dots,r\}, \\ P'^n\left[\zeta_k^{n\ell j} = 1\right] = 1 - P'^n\left[\zeta_k^{n\ell j} = 0\right] &= \frac{T}{n}, \quad \ell,j \in \{1,\dots,m\}, \end{aligned}$$

and define for $n \in \mathbb{N}$, $i \in \{1, \dots, r\}$ and $\ell, j \in \{1, \dots, m\}$ the stochastic processes

$$L_t^n = \left[\frac{nt}{T}\right]\frac{T}{n}, \quad W_t^{ni} = \sum_{k=1}^{\left[\frac{nt}{T}\right]} \xi_k^{ni}, \quad N_t^{n\ell j} = \sum_{k=1}^{\left[\frac{nt}{T}\right]} \zeta_k^{n\ell j}, \quad t \in [0,T]. \tag{5.10}$$

If we write $W^n = (W^{n1}, \dots, W^{nr})$ and $N^n = (N^{n11}, \dots, N^{nmm})$ and let $\mathbb{F}^n = \mathbb{F}^{(W^n,N^n)}(P'^n)$, the P'^n-augmentation of the (right-continuous) filtration generated by W^n and N^n, then obviously (L^n, W^n, N^n) is an $(\mathbb{F}^n, P'^n)$-semimartingale with càdlàg paths, independent increments and values in $\mathbb{R} \times \mathbb{R}^r \times \mathbb{R}^{m\times m}$. Note that L^n is a discretization of time, whereas W^n and N^n are binomial random walks which are constructed to converge to W and N, respectively. To ease notation we define $n(t) := \left[\frac{nt}{T}\right]$. We show convergence of (L^n, W^n, N^n) with the help of Theorem 1.46. To that end let $L_t := t$ and recall that W was defined to be an r-dimensional P'-standard Brownian motion and N a multivariate standard Poisson process under P'.

Proposition 5.2 *Let L^n, W^n and N^n be as above. Then*

$$\mathcal{L}\left(L^n, W^n, N^n \,\middle|\, P'^n\right) \xrightarrow{w} \mathcal{L}(L, W, N | P')$$

as $n \to \infty$.

Proof. Since the coordinates of (L^n, W^n, N^n) are independent under P'^n it suffices to show convergence in distribution of the single coordinates, i.e. $L^n \xrightarrow{\mathcal{L}} L$, $W^{ni} \xrightarrow{\mathcal{L}} W^i$ and $N^{n\ell j} \xrightarrow{\mathcal{L}} N^{\ell j}$ for $i \in \{1, \dots, r\}$ and $\ell, j \in \{1, \dots, m\}$; cf. Billingsley (1968), Theorem 3.2.

Now L^n as well as L is deterministic with $L_t^n \overset{n\to\infty}{\longrightarrow} t$ for all $t \in [0,T]$, so there is nothing to show here. Convergence of W^{ni} to W^i for $i \in \{1, \dots, r\}$ is immediate by Donsker's theorem, so it remains to prove convergence of $N^{n\ell j}$ to $N^{\ell j}$ for $\ell, j \in \{1, \dots, m\}$. But $N^{\ell j}$ for $\ell, j \in \{1, \dots, m\}$ is a standard Poisson process, so $N^{n\ell j} \xrightarrow{\mathcal{L}} N^{\ell j}$ is just a functional version of the Poisson limit law, which should be well known (cf. Rachev and Rüschendorf (1994), Section 2.5). For the sake of completeness we give the details, using Theorem 1.46.

First we need the characteristics of $N^{n\ell j}$ and $N^{\ell j}$ for $\ell, j \in \{1, \dots, m\}$ under P'^n and P', respectively. Both $N^{n\ell j}$ and $N^{\ell j}$ have bounded jumps, and are therefore special semimartingales, so we can choose the "truncation function" $h(x) = x$, and we get the canonical decomposition under P' as

$$N_t^{\ell j} = (N_t^{\ell j} - t) + t.$$

Therefore the P'-characteristics of $N^{\ell j}$ are

$$\begin{aligned} B_t^{\ell j} &= t, \\ C_t^{\ell j} &= 0, \\ \nu^{\ell j}(dt, dx) &= dt\ \delta_{\{1\}}(dx), \end{aligned}$$

where $\delta_{\{1\}}$ denotes the Dirac measure in the point 1. (Note that $(B^{\ell j}, C^{\ell j}, \nu^{\ell j})$ are not the ℓ-j-th coordinates of the characteristics of N!) With this and Proposition 1.25 we get the modified second characteristic $\tilde{C}^{\ell j}$ by

$$\tilde{C}_t^{\ell j} = x^2 * \nu_t^{\ell j} = t.$$

Concerning the P'^n-characteristics of $N^{n\ell j}$ we see that the canonical decomposition of $N^{n\ell j}$ is given by

$$N_t^{n\ell j} = \left(\sum_{k=1}^{n(t)} \left(\zeta_k^{n\ell j} - E_{P'^n}\left[\zeta_1^{n\ell j}\right] \right) \right) + n(t) E_{P'^n}\left[\zeta_1^{n\ell j}\right],$$

so that for $g \geq 0$ measurable the P'^n-characteristics of $N^{n\ell j}$ are

$$\begin{aligned} B_t^{n\ell j} &= n(t)\frac{T}{n}, \\ C_t^{n\ell j} &= 0, \\ g * \nu_t^{n\ell j} &= n(t) E_{P'^n}\left[g\left(\zeta_1^{n\ell j}\right) \mathbb{1}_{\left\{\zeta_1^{n\ell j} \neq 0\right\}} \right] = n(t) g(1) \frac{T}{n}. \end{aligned}$$

For this result, especially the form of the third characteristic, also cf. Jacod and Shiryaev (1987), Theorem II.3.11. Finally the modified second characteristic $\tilde{C}^{n\ell j}$ of $N^{n\ell j}$ under P'^n is then given by (see again Proposition 1.25)

$$\tilde{C}_t^{n\ell j} = x^2 * \nu_t^{n\ell j} - \sum_{s \leq t} (\Delta B_s^{n\ell j})^2 = n(t) \operatorname{Var}_{P'^n}\left[\zeta_1^{n\ell j}\right] = n(t) \frac{T}{n}\left(1 - \frac{T}{n}\right).$$

Now convergence to 0 of $\sup_{0 \leq s \leq T} |B_s^{n\ell j} - B_s^{\ell j}|$ and $|\tilde{C}_t^{n\ell j} - \tilde{C}_t^{\ell j}|$ for all $t \in [0, T]$ is immediate, whereas for $g \in \mathcal{C}(\mathbb{R})$ we have

$$g * \nu_t^{n\ell j} = n(t) E_{P'^n}\left[g\left(\zeta_1^{n\ell j}\right) \mathbb{1}_{\left\{\zeta_1^{n\ell j} \neq 0\right\}} \right] = n(t) \frac{T}{n}\, g(1) \stackrel{n \to \infty}{\longrightarrow} t g(1) = g * \nu_t^{\ell j},$$

which ends the proof. □

Proposition 5.3 *The sequence $(L^n, W^n, N^n)_{n \in \mathbb{N}}$ is good with respect to P'^n, P'.*

Proof. For all $k \in \{1, \ldots, n\}$ we define the $1 + r + m^2$-dimensional random vector $\bar{\xi}_k$ by $\bar{\xi}_k^{n0} := \frac{T}{n}$, $\bar{\xi}_k^{ni} = \xi_k^{ni}$ for $i \in \{1, \ldots, r\}$, and $\bar{\xi}_k^{n,r+1} := \zeta_k^{n11}, \ldots, \bar{\xi}_k^{n,r+m^2} = \zeta_k^{nmm}$. Then, using Example 1.51, we need to show that $E_{P'^n}\left[\bar{\xi}_1^n\right] = \mathcal{O}\left(\frac{1}{n}\right)$ and $\operatorname{Cov}_{P'^n}\left(\bar{\xi}_1^{ni}, \bar{\xi}_1^{nj}\right) = \mathcal{O}\left(\frac{1}{n}\right)$ for $i, j \in \{0, \ldots, r + m^2\}$.

Now $\left|E_{P'^n}\left[\bar{\xi}_1^n\right]\right|^2 = \left(\frac{T}{n}\right)^2 + m^2\left(\frac{T}{n}\right)^2 = (1 + m^2)\left(\frac{T}{n}\right)^2$, so that $n\left|E_{P'^n}\left[\bar{\xi}_1^n\right]\right| = \sqrt{1 + m^2}\, T$, which is clearly bounded.

Concerning the covariances we see that $\mathrm{Cov}_{P'^n}\left(\bar{\xi}_1^{ni}, \bar{\xi}_1^{nj}\right) = 0$ for $i \neq j$ or $i = 0$ or $j = 0$, whereas in the other cases we have

$$\mathrm{Var}_{P'^n}\left[\bar{\xi}_1^{ni}\right] = \begin{cases} \frac{T}{n} & \text{if } i = j \in \{1, \ldots, r\} \\ \frac{T}{n}\left(1 - \frac{T}{n}\right) & \text{if } i = j \in \{r+1, \ldots, r+m^2\} \end{cases}$$

so that $n\,\mathrm{Cov}_{P'^n}\left(\bar{\xi}_1^{ni}, \bar{\xi}_1^{nj}\right)$ is bounded. □

We now define the discrete-time processes S^n and η^n as solutions of the difference equations

$$dS_t^n = \Gamma(t, S_{t-}^n, \eta_{t-}^n)\, dL_t^n + \Sigma(t, S_{t-}^n, \eta_{t-}^n)\, dW_t^n, \tag{5.11}$$

$$d\eta_t^n = \sum_{\ell,j=1}^{m} (j-\ell) \mathbb{1}_{\{\eta_{t-}^n = \ell\}}\, dN_t^{n\ell j}, \tag{5.12}$$

with $S_0^n = S_0 \in \mathbb{R}^d$ and $\eta_0^n = \eta_0 \in \{1, \ldots, m\}$.

Note that in contrast to the continuous-time case, $N^{n\ell j_1}$ and $N^{n\ell j_2}$ may jump simultaneously for $j_1 \neq j_2$ so that it may happen that η^n jumps out of the set $\{1, \ldots, m\}$. However if $\eta_{t_0}^n \notin \{1, \ldots, m\}$ for some $t_0 > 0$, then by construction $d\eta_t^n = 0$ for all $t > t_0$, so that η stays constant once it has left the set $\{1, \ldots, m\}$. We therefore identify all points outside of $\{1, \ldots, m\}$ with the cemetery point π, and we write $\eta_{t_k}^n = \pi$ to mean that $\eta_{t_k}^n \notin \{1, \ldots, m\}$. Note that the range of η^n is therefore $I_m = \{\pi\} \cup \{1, \ldots, m\}$. In the sequel we shall view I_m as a subset of $\mathbb{Z}$.

Also notice that due to the distribution of $\zeta_k^{n\ell j}$ the probability that $N^{n\ell j_1}$ and $N^{n\ell j_2}$ jump simultaneously is small and vanishes in the limit as $n \to \infty$. For a more detailed analysis of the processes η^n see Section 5.4, where it is also shown that the probability that η^n leaves $\{1, \ldots, m\}$ vanishes as $n \to \infty$.

Proposition 5.4 *We have*

$$\mathcal{L}(L^n, W^n, N^n, \eta^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, \eta | P')$$

as $n \to \infty$.

Proof. In order to apply Theorem 1.53 we need to adjust the dimension of the processes $N = (N^{\ell j})_{\ell, j \in \{1, \ldots, m\}}$ and $N^n = (N^{n\ell j})_{\ell, j \in \{1, \ldots, m\}}$. So we associate to any $k \in \{1, \ldots, m^2\}$ the numbers ℓ and j by $\ell = \left[\frac{k-1}{m}\right] + 1$ and $j = 1 + ((k-1) \bmod m)$, so that $\ell, j \in \{1, \ldots, m\}$, and we define the $\mathbb{Z}^{m^2}$-valued processes $\bar{N}$ by $\bar{N}^k = N^{\ell j}$ and $\bar{N}^n$ by $\bar{N}^{nk} = N^{n\ell j}$. Then for $f \colon \mathbb{Z} \to \mathbb{Z}^{m^2}$, given by $f^k(y) = (j - \ell)\mathbb{1}_{\{y = \ell\}}$, $k \in \{1, \ldots, m^2\}$, the stochastic differential equations (5.2) and (5.12) may be written as

$$\begin{aligned} d\eta_t &= f(\eta_{t-})\, d\bar{N}_t, \quad \eta_0 \in \{1, \ldots, m\}, \\ d\eta_t^n &= f(\eta_{t-}^n)\, d\bar{N}_t^n, \quad \eta_0^n = \eta_0 \in \{1, \ldots, m\}. \end{aligned}$$

Now $(L^n, W^n, N^n) \xrightarrow{\mathcal{L}} (L, W, N)$ by Proposition 5.2, and since (L, W) is a continuous semimartingale, the result follows from Theorem 1.53. $\square$

In analogy to the continuous-time case we define the sequence of jump times of η^n by

$$\begin{cases} \tau_0^n &= 0, \\ \tau_k^n &= \inf\left\{t > \tau_{k-1}^n \mid |\Delta\eta_t^n| > \frac{1}{2}\right\} \wedge T. \end{cases} \tag{5.13}$$

Now let $\ell \in I_m$ and let $S^{(k),n}$, $S^{(k,\ell),n}$ and $S^{(k),\ell,n}$ be the discrete-time processes which correspond to $S^{(k)}$, $S^{(k,\ell)}$ and $S^{(k),\ell}$, i.e. $S^{(0),n} \equiv S_0$, and $S^{(k+1),n}$ is the solution of the difference equation

$$\begin{aligned} S_t^{(k+1),n} &= S_{t\wedge\tau_k^n}^{(k),n} + \int_0^t \mathbb{1}_{]\!]\tau_k^n,\tau_{k+1}^n]\!]}\, \Gamma(s, S_{s-}^{(k+1),n}, \eta_{\tau_k^n}^n)\, dL_s^n \\ &\quad + \int_0^t \mathbb{1}_{]\!]\tau_k^n,\tau_{k+1}^n]\!]}\, \Sigma(s, S_{s-}^{(k+1),n}, \eta_{\tau_k^n}^n)\, dW_s^n \end{aligned}$$

for $k \geq 0$. As above we define $S^{(k,\ell),n}$ recursively by setting $S_t^{(0,\ell),n} = S_0$ and for $\ell \in I_m$

$$\begin{aligned} S_t^{(k+1,\ell),n} &= S_{t\wedge\tau_k^n}^{(k),n} + \int_0^t \mathbb{1}_{]\!]\tau_k^n,\tau_{k+1}^n]\!]}\, \Gamma(s, S_{s-}^{(k+1,\ell),n}, \ell)\, dL_s^n \\ &\quad + \int_0^t \mathbb{1}_{]\!]\tau_k^n,\tau_{k+1}^n]\!]}\, \Sigma(s, S_{s-}^{(k+1,\ell),n}, \ell)\, dW_s^n. \end{aligned} \tag{5.14}$$

For the above $S^{(k),n}$ we define $S^{(k),\ell,n}$ for $\ell \in I_m$ by

$$S_t^{(k),\ell,n} = \int_0^t \Gamma(s, S_{s-}^{(k),n}, \ell)\, dL_s^n + \int_0^t \Sigma(s, S_{s-}^{(k),n}, \ell)\, dW_s^n, \quad t \in [0, T],$$

and with a similar argument as in Lemma 5.1 we can write the stochastic differential equation (5.14) as

$$\begin{aligned} S_t^{(k+1,\ell),n} &= S_{t\wedge\tau_k^n}^{(k),n} - S_{t\wedge\tau_k^n}^{(k),\ell,n} + \int_0^{t\wedge\tau_{k+1}^n} \Gamma(s, S_{s-}^{(k+1,\ell),n}, \ell)\, dL_s^n \\ &\quad + \int_0^{t\wedge\tau_{k+1}^n} \Sigma(s, S_{s-}^{(k+1,\ell),n}, \ell)\, dW_s^n. \end{aligned} \tag{5.15}$$

Like in the continuous-time case we have

$$(S^n)^{\tau_{k+1}^n} = S^{(k+1),n} = \sum_{\ell\in I_m} S^{(k+1,\ell),n}\, \mathbb{1}_{\{\eta_{\tau_k^n}^n = \ell\}}.$$

Proposition 5.5 *We have for all $k \in \mathbb{N}$*

$$\mathcal{L}(L^n, W^n, N^n, \eta^n, S^{(k),n} | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, \eta, S^{(k)} | P')$$

as $n \to \infty$.

Proof. We use induction over k. For $k = 0$ we have $S^{(0),n} \equiv S_0 \equiv S^{(0)}$, so the claim is obvious because $(L^n, W^n, N^n, \eta^n) \xrightarrow{\mathcal{L}} (L, W, N, \eta)$ by Proposition 5.4. So suppose

$$(L^n, W^n, N^n, \eta^n, S^{(k),n}) \xrightarrow{\mathcal{L}} (L, W, N, \eta, S^{(k)})$$

for some $k \geq 0$. Then because $(L^n, W^n, N^n)_{n \in \mathbb{N}}$ is good by Proposition 5.3 we have

$$\Big(L^n, W^n, N^n, \eta^n, S^{(k),n}, (S^{(k),\ell,n})_{\ell \in I_m}\Big) \xrightarrow{\mathcal{L}} \Big(L, W, N, \eta, S^{(k)}, (S^{(k),\ell})_{\ell \in I_m}\Big)$$

by construction of $S^{(k),\ell,n}$ and $S^{(k),\ell}$ and the definition of goodness (cf. Definition 1.48). Then since $S^{(k),\ell}$ is continuous we have convergence of the corresponding stopped processes $(S^{(k),\ell,n})^{\tau_k^n}$, $\ell \in I_m$, by Proposition B.6, and thus

$$\Big(L^n, W^n, N^n, \eta^n, (J^{\ell,n}, S^{(k),\ell,n})_{\ell \in I_m}\Big) \xrightarrow{\mathcal{L}} \Big(L, W, N, \eta, (J^\ell, S^{(k),\ell})_{\ell \in I_m}\Big),$$

where $J_t^{\ell,n} = S_{t \wedge \tau_k^n}^{(k),n} - S_{t \wedge \tau_k^n}^{(k),\ell,n}$ and analogously $J_t^\ell = S_{t \wedge \tau_k}^{(k)} - S_{t \wedge \tau_k}^{(k),\ell}$. Due to the representations (5.7) and (5.15) of $S^{(k+1,\ell)}$ and $S^{(k+1,\ell),n}$ we have convergence of

$$\Big(L^n, W^n, N^n, \eta^n, (S^{(k+1,\ell),n})_{\ell \in I_m}\Big) \xrightarrow{\mathcal{L}} \Big(L, W, N, \eta, (S^{(k+1,\ell)})_{\ell \in I_m}\Big)$$

as a result of Kurtz and Protter (1996), Theorem 8.6. Finally by Proposition B.7 and the continuous mapping theorem we get

$$(L^n, W^n, N^n, \eta^n, S^{(k+1),n}) \xrightarrow{\mathcal{L}} (L, W, N, \eta, S^{(k+1)}),$$

which finishes the proof. □

Theorem 5.6 *Let W be an r-dimensional standard Brownian motion and N a multivariate standard Poisson process, independent of W under P'. Let (S, η) be given by (5.1) and (5.2) and let $L_t = t$. Then for $(L^n, W^n, N^n, \eta^n, S^n)$ given by (5.10), (5.11) and (5.12) we have*

$$\mathcal{L}(L^n, W^n, N^n, \eta^n, S^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, \eta, S | P')$$

as $n \to \infty$.

Proof. Abbreviate $X^n = (L^n, W^n, N^n, \eta^n, S^n)$, $X^{(k),n} = (L^n, W^n, N^n, \eta^n, S^{(k),n})$, and let $X, X^{(k)}$ be the corresponding continuous-time processes. By Proposition 5.5 we have that $X^{(k),n} \xrightarrow{\mathcal{L}} X^{(k)}$ for all $k \in \mathbb{N}$. We first show tightness of $(X^n)_{n \in \mathbb{N}}$ with the help of Theorem 1.43 and then identify X as the only possible limit process. In our situation of a finite time horizon $[0, T]$ it suffices to show the conditions of Theorem 1.43 for $N = T$ only. To that end notice that

$$\lim_{k \to \infty} P'^n[\tau_k^n = T] = 1, \quad \lim_{k \to \infty} P'[\tau_k = T] = 1 \tag{5.16}$$

for all $n \in \mathbb{N}$, since η^n and η have only finitely many jumps in $[0,T]$, and that

$$\lim_{n\to\infty} P'^n[\tau_k^n = T] = P'[\tau_k = T], \tag{5.17}$$

for all $k \in \mathbb{N}$, since $\eta^n \xrightarrow{\mathcal{L}} \eta$ and therefore $\tau_k^n \xrightarrow{\mathcal{L}} \tau_k$ by Proposition B.5. Furthermore on $\{\tau_k^n = T\}$ we have $X^n = X^{(k),n}$, therefore we can write for arbitrary n, K and k

$$\begin{aligned} P'^n\left[\sup_{0\le t\le T}|X_t^n| \le K\right] &\ge P'^n\left[\sup_{0\le t\le T}|X_t^n| \le K, \tau_k^n = T\right] \\ &= P'^n\left[\sup_{0\le t\le T}|X_t^{(k),n}| \le K, \tau_k^n = T\right] \\ &\ge P'^n\left[\sup_{0\le t\le T}|X_t^{(k),n}| \le K\right] + P'^n\left[\tau_k^n = T\right] - 1. \end{aligned}$$

Now let $\varepsilon > 0$ and choose k_0 large enough such that $P'[\tau_{k_0} = T] \ge 1 - \frac{\varepsilon}{4}$ by (5.16). Then by (5.17) we can choose $n_1 \in \mathbb{N}$ such that $P'^n[\tau_{k_0}^n = T] \ge 1 - \frac{\varepsilon}{2}$ for all $n \ge n_1$. Now $(X^{(k_0),n})_{n\in\mathbb{N}}$ is tight, so we can choose $K > 0$ and $n_2 \in \mathbb{N}$ so that $P'^n\left[\sup_{0\le t\le T}\left|X_t^{(k_0),n}\right| \le K\right] \ge 1 - \frac{\varepsilon}{2}$ for all $n \ge n_2$. Then for all $n \ge n_0 := n_1 \vee n_2$ we have

$$P'^n\left[\sup_{0\le t\le T}|X_t^n| \le K\right] \ge 1 - \varepsilon,$$

which yields condition (i) of Theorem 1.43. Condition (ii) is shown in the same manner: Let $\varepsilon, \delta > 0$ and choose k_0 and n_1 as above. By the tightness of $(X^{(k_0),n})_{n\in\mathbb{N}}$ we can choose $n_2 \in \mathbb{N}$ and $\vartheta > 0$ such that for all $n \ge n_2$ we have $P'^n\left[w'_N(X^{(k_0),n}, \vartheta) \le \delta\right] \ge 1 - \frac{\varepsilon}{2}$, so that with the same argument as above we have

$$P'^n\left[w'_N(X^n, \vartheta) \le \delta\right] \ge 1 - \varepsilon,$$

which finally shows that $(X^n)_{n\in\mathbb{N}}$ is tight.

It remains to identify X as the only possible limit of X^n. So let X^{ℓ_n} be a weakly convergent subsequence with limit process Y. Then on the one hand we have that $X^{(k),\ell_n} \xrightarrow{\mathcal{L}} X^{(k)}$ for all $k \in \mathbb{N}$ by Proposition 5.5, and on the other hand we have $X^{(k),\ell_n} = \varphi_k(X^{\ell_n})$, where φ_k is the Skorokhod-continuous function from Proposition B.6, which stops the last component of a càdlàg function at time τ_k. Since the last component S of the limit process X is continuous, we have by the continuous mapping theorem

$$X^{(k),\ell_n} = \varphi_k(X^{\ell_n}) \xrightarrow{\mathcal{L}} \varphi_k(Y) =: Y^{(k)}$$

for all $k \in \mathbb{N}$ and therefore $Y^{(k)} = X^{(k)}$ for all $k \in \mathbb{N}$. Since $P'[\tau_k = T] \uparrow 1$ as $k \to \infty$, this yields $Y = X$. $\square$

5.3 Convergence under Mutual Dependences

In order to approximate the model given by (5.1), (5.2) and (5.3) (i.e. where $N_t^{\ell j}$ has $(P,\mathbb{F})$-intensity $\lambda^{\ell j}(t,S_t)$ for $\ell,j\in\{1,\dots,m\}$), we need convergence of (S^n,η^n) not under P'^n,P' but under some measures P^n,P such that the mutual dependences of S and η hold as specified in (5.3). To that end we approximate the density process Z of P with respect to P' in the following way. Recall that for $t\in[0,T]$ and $n\in\mathbb{N}$ we have defined $n(t)=\left[\frac{nt}{T}\right]$. For $k\in\{0,\dots,n\}$ and fixed $n\in\mathbb{N}$ we furthermore define $t_k=k\frac{T}{n}$, and we denote $t_k-t_{k-1}=\frac{T}{n}$ by Δt_k. For $n\in\mathbb{N}$ and $\ell,j\in\{1,\dots,m\}$ let the process $M^{n\ell j}$ be given by

$$M_t^{n\ell j}=\int_0^t\left(\lambda^{\ell j}\left(s,S_{s-}^n\right)-1\right)\left(dN_s^{n\ell j}-dL_s^n\right)=\sum_{k=1}^{n(t)}\left(\lambda^{\ell j}\left(t_k,S_{t_{k-1}}^n\right)-1\right)\left(\zeta_k^{n\ell j}-\Delta t_k\right),$$

$t\in[0,T]$, and define the processes $Z^{n\ell j}$ by

$$Z_t^{n\ell j}=\mathcal{E}\left(M^{n\ell j}\right)_t=\prod_{k=1}^{n(t)}\left(1+\Delta M_{t_k}^{n\ell j}\right),\quad t\in[0,T],$$

where $\Delta M_{t_k}^{n\ell j}$ stands for $M_{t_k}^{n\ell j}-M_{t_{k-1}}^{n\ell j}$. Finally let

$$Z_t^n=\prod_{\ell,j=1}^m Z_t^{n\ell j},\quad t\in[0,T]. \tag{5.18}$$

Proposition 5.7 *Let Z^n be defined as above. Then for n sufficiently large Z^n is a nonnegative P'^n-martingale with $E_{P'^n}[Z_T^n]=1$.*

Proof. Since $0\le\lambda^{\ell j}\le K$ for some constant K, we have

$$\Delta M_{t_k}^{n\ell j}=(\lambda^{\ell j}(t_k,S_{t_{k-1}}^n)-1)\left(\zeta_k^{n\ell j}-\Delta t_k\right)\ge -1$$

P'-a.s. for n sufficiently large, so that $Z^{n\ell j}\ge 0$ P'-a.s. for all $\ell,j\in\{1,\dots,m\}$ and n sufficiently large, and thus $Z^n\ge 0$ P'-a.s. for n sufficiently large. Furthermore we have for all $k\in\{0,\dots,n\}$ by the independence of the $\zeta_k^{n\ell j}$ under P'^n

$$\begin{aligned}E_{P'^n}\left[\frac{Z_{t_k}^n}{Z_{t_{k-1}}^n}\middle|\mathcal{F}_{t_{k-1}}^n\right] &= E_{P'^n}\left[\prod_{\ell,j=1}^m\left(1+\Delta M_{t_k}^{n\ell j}\right)\middle|\mathcal{F}_{t_{k-1}}^n\right]\\ &= E_{P'^n}\left[\prod_{\ell,j=1}^m\left(1+\left(\lambda^{\ell j}\left(t_k,x\right)-1\right)\left(\zeta_k^{n\ell j}-\Delta t_k\right)\right)\right]\Bigg|_{x=S_{t_{k-1}}^n}\\ &= \prod_{\ell,j=1}^m\left(1+\left(\lambda^{\ell j}\left(t_k,S_{t_{k-1}}^n\right)-1\right)E_{P'^n}\left[\left(\zeta_k^{n\ell j}-\Delta t_k\right)\right]\right)\\ &= 1,\end{aligned}$$

so that Z^n is a martingale under P'^n. Since $Z_0^n = 1$, this implies $E_{P'^n}[Z_T^n] = 1$. $\square$

Now we can define a measure $P^n \ll P'^n$ on $\mathcal{F}_T^n$ by

$$dP^n = Z_T^n \, dP'^n, \tag{5.19}$$

and Z^n is obviously the density process of P^n with respect to P'^n.

Lemma 5.8 *Let $P^n \ll P'^n$ be given as above. Then*

$$N_t^{n\ell j} - \int_0^t \bar{\lambda}^{n\ell j}(s, S_{s-}^n) \, dL_s^n$$

is a P^n-martingale for $\bar{\lambda}^{n\ell j}(t,x) = \lambda^{\ell j}(t,x) + \frac{T}{n}\left(1 - \lambda^{\ell j}(t,x)\right)$.

Proof. It suffices to show that $E_{P^n}\left[\Delta N_k^{n\ell j} \middle| \mathcal{F}_{t_{k-1}}^n\right] = \bar{\lambda}^{n\ell j}(t_k, S_{t_{k-1}}^n) \, \Delta t_k$. By the Bayes formula and the independence of the $\zeta_k^{n\ell j}$ under P'^n we get

$$\begin{aligned}
E_{P^n}\left[\Delta N_k^{n\ell j} \middle| \mathcal{F}_{t_{k-1}}^n\right] &= E_{P'^n}\left[\zeta_k^{n\ell j}\left(1 + \left(\lambda^{\ell j}(t_k, S_{t_{k-1}}^n) - 1\right)\left(\zeta_k^{n\ell j} - \frac{T}{n}\right)\right) \middle| \mathcal{F}_{t_{k-1}}^n\right] \\
&= E_{P'^n}\left[\zeta_k^{n\ell j}\left(1 + \left(\lambda^{\ell j}(t_k, x) - 1\right)\left(\zeta_k^{n\ell j} - \frac{T}{n}\right)\right)\right]\Bigg|_{x = S_{t_{k-1}}^n} \\
&= \left(1 + \left(\lambda^{\ell j}(t_k, S_{t_{k-1}}^n) - 1\right)\left(1 - \frac{T}{n}\right)\right)\frac{T}{n} \\
&= \bar{\lambda}^{n\ell j}(t_k, S_{t_{k-1}}^n) \, \Delta t_k
\end{aligned}$$

by the distribution of $\zeta_k^{n\ell j}$ under P'^n. $\square$

Remark 5.9 Of course for the purpose of approximating the density process Z of P with respect to P' from (5.9), the construction of a sequence of density processes Z^n is by no means unique. However there are some aspects where one needs to be careful.

a) First note that the construction of Z^n is not exactly analogous to the construction of Z. But if we define

$$\tilde{Z}^n = \mathcal{E}\left(\sum_{\ell,j=1}^m \int \left(\lambda^{\ell j}\left(s, S_{s-}^n\right) - 1\right)\left(dN_s^{nj} - dL_s^n\right)\right),$$

then $\tilde{Z}^n$ may become negative since the jumps of the exponent in the stochastic exponential are not bounded from below by -1 any more. In fact, the processes $N^{n\ell j}$ may jump simultaneously so that in the worst case the exponent may jump by $-m^2$. The proof of the following Proposition 5.10 will make clear that this phenomenon vanishes in the limit.

b) By Lemma 5.8 the process $N^{n\ell j} - \int \bar{\lambda}^{n\ell j}(s, S_{s-}^n) \, dL^n$ is a P^n-martingale, so that it is justified to say that $dN_t^{n\ell j}$ has P^n-"intensity" $\bar{\lambda}^{n\ell j}(t, S_t^n) \, dL_t^n$, in analogy to the P-intensities

$\lambda^{\ell j}(t, S_t)\ dt$ of $dN_t^{\ell j}$. As $n \to \infty$, we clearly have $\bar{\lambda}^{n\ell j} \to \lambda^{\ell j}$. If one wants to obtain the original "intensities" $\lambda^{\ell j}(t, S_t^n)\ dL_t^n$ under some measure $\hat{P}^n$ one needs to define the density process $\hat{Z}^n$ by

$$\hat{Z}_t^n = \prod_{\ell,j=1}^{m} \mathcal{E}\left(\hat{M}^{n\ell j}\right)_t$$

with

$$\hat{M}_t^{n\ell j} = \frac{1}{1 - \frac{T}{n}} \int_0^t \left(\lambda^{\ell j}\left(s, S_{s-}^n\right) - 1\right)\left(dN_s^{n\ell j} - dL_s^n\right).$$

In this case, with $\hat{P}^n$ defined via $\hat{Z}_T^n$,

$$N_t^{n\ell j} - \int_0^t \lambda^{\ell j}(s, S_{s-}^n)\ dL_s^n = N_t^{n\ell j} - \sum_{k=1}^{n(t)} \lambda^{\ell j}\left(t_k, S_{t_{k-1}}^n\right)\ \Delta t_k$$

is a $\hat{P}^n$-martingale, so that $dN_t^{n\ell j}$ has $\hat{P}^n$-"intensities" $\lambda^{\ell j}(t, S_t^n)\ dL_t^n$. We have chosen to construct P^n via Z^n and M^n as above since then the convergence of the density processes Z^n is easier to prove, and the difference of intensities vanishes in the limit anyway. $\diamond$

Proposition 5.10 *For all $n \in \mathbb{N}$, let Z^n be defined by (5.18) and let Z be the density process of P with respect to P' from (5.9). Then*

$$\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, Z^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta, Z | P').$$

Proof. Recall that $\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta | P')$ by Theorem 5.6. We define for $n \in \mathbb{N}$ the m^2-dimensional processes U^n by $U_t^{n\ell j} = \lambda^{\ell j}(t, S_t^n) - 1$ and U by $U_t^{\ell j} = \lambda^{\ell j}(t, S_t) - 1$, $\ell, j \in \{1, \ldots, m\}$. Then

$$\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, U^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta, U | P')$$

by Proposition B.4 and the continuous mapping theorem. Furthermore

$$M^{n\ell j} = \int U_-^{n\ell j}\ (dN^{n\ell j} - dL^n), \quad \ell, j \in \{1, \ldots, m\},$$

so goodness of $(L^n, W^n, N^n)_{n\in\mathbb{N}}$ under P'^n, P' (cf. Proposition 5.3) implies

$$\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, U^n, M^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta, U, M | P'),$$

where M^n and M are $\mathbb{R}^{m^2}$-valued with $M^{\ell j} = \int U_-^{\ell j}\left(dN^{\ell j} - ds\right)$. However $(M^n)_{n\in\mathbb{N}}$ is also good by Proposition 1.49, so for $Z^{n\ell j} = \mathcal{E}(M^{n\ell j})$, the solution of $dZ^{n\ell j} = Z_-^{n\ell j}\ dM^{n\ell j}$, we get

$$\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, (Z^{n\ell j})_{\ell,j} | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta, (Z^{\ell j})_{\ell,j} | P')$$

by Theorem 1.52, where $Z^{\ell j} = \mathcal{E}(M^{\ell j})$. Now the mapping $\varphi\colon \mathbb{D}(\mathbb{R}^{m^2}) \to \mathbb{D}(\mathbb{R})$, given by $\varphi(\alpha)(s) = \prod_{\ell,j=1}^m \alpha^{\ell j}(s)$, is continuous for the Skorokhod topology by Proposition B.4, so that by the continuous mapping theorem

$$\mathcal{L}\left(L^n, W^n, N^n, S^n, \eta^n, Z^n | P'^n\right) \xrightarrow{w} \mathcal{L}\left(L, W, N, S, \eta, \prod_{\ell,j=1}^m Z^{\ell j} \middle| P'\right),$$

and since the M^j are (compensated) pure jump processes which never jump simultaneously because the driving processes are independent Poisson processes under P', we have $[M^j, M^k] = 0$ for $j \neq k$ so that

$$\prod_{\ell,j=1}^m Z^{\ell j} = \prod_{\ell,j=1}^m \mathcal{E}(M^{\ell j}) = \mathcal{E}\left(\sum_{\ell,j=1}^m M^{\ell j}\right) = Z$$

by Yor's formula; cf. Protter (1990), Theorem II.37. This yields the result. □

Proposition 5.11 *If* $0 \leq \lambda^{\ell j}(t,x) \leq K < \infty$ *for all* $t \in [0,T]$*,* $x \in \mathbb{R}^d$ *and* $\ell, j \in \{1, \ldots, m\}$*, then* $(P^n)_{n\in\mathbb{N}} \overset{\text{loc}}{\lhd} (P'^n)_{n\in\mathbb{N}}$.

Proof. By Lemma 1.55 local contiguity of $(P^n)_{n\in\mathbb{N}}$ with respect to $(P'^n)_{n\in\mathbb{N}}$ is equivalent to

$$\lim_{\alpha\downarrow 0} \liminf_{n\to\infty} H(\alpha; P'^n_t, P^n_t) = 1$$

for all $t \geq 0$, where H is the Hellinger integral of order α, and since $P^n \overset{\text{loc}}{\ll} P'^n$ for all $n \in \mathbb{N}$, the Hellinger integral becomes $H(\alpha; P'^n_t, P^n_t) = E_{P'^n}\left[(Z^n_t)^{1-\alpha}\right]$. So we need to show

$$\lim_{\alpha\downarrow 0} \liminf_{n\to\infty} E_{P'^n}\left[(Z^n_t)^{1-\alpha}\right] = 1$$

for all t.

To that end we remark that by Jensen's inequality we have

$$E_{P'^n}\left[(Z^n_t)^{1-\alpha}\right] \leq 1 \tag{5.20}$$

for all $t \geq 0$, $n \in \mathbb{N}$, $\alpha \in (0,1)$. On the other hand we have by conditioning on $\mathcal{F}^n_{t_{k-1}}$ and the independence of the $(\zeta^{n\ell j}_k)_{\ell,j,k}$ under P'^n

$$E_{P'^n}\left[(Z^n_t)^{1-\alpha}\right] = E_{P'^n}\left[\prod_{\ell,j=1}^m \prod_{k=1}^{n(t)} \left(1 + \Delta M^{n\ell j}_{t_k}\right)^{1-\alpha}\right] = E_{P'^n}\left[\prod_{\ell,j=1}^m \prod_{k=1}^{n(t)} \varphi^{\ell j}_{k,n}(S^n_{t_{k-1}})\right], \tag{5.21}$$

with

$$\varphi^{\ell j}_{k,n}(x) = E_{P'^n}\left[\left(1 + \left(\lambda^{\ell j}(t_k, x) - 1\right)\left(\zeta^{n\ell j}_k - \Delta t_k\right)\right)^{1-\alpha}\right].$$

We define $f_\alpha(y) = y^{1-\alpha} - y + \alpha(y-1)$ for $y \geq 0$, and we claim that for n sufficiently large we have

$$\varphi^{\ell j}_{k,n}(x) \geq 1 + \frac{T f_\alpha(\bar\lambda) + \bar R_n}{n} \tag{5.22}$$

uniformly for all x, k, ℓ, j, for some $0 \leq \bar\lambda < \infty$, and $\bar R_n = \mathcal{O}\left(\frac{1}{n}\right)$.

Indeed by the distribution of $\zeta^{n\ell j}_k$ under P'^n we have for fixed $0 \leq y \leq K$ (recall that $\Delta t_k = t_k - t_{k-1} = \frac{T}{n}$)

$$\begin{aligned} & E_{P'^n}\left[\left(1 + (y-1)\left(\zeta^{n\ell j}_k - \Delta t_k\right)\right)^{1-\alpha}\right] \\ = \; & \frac{T}{n}\left(1 + (y-1)\left(1 - \frac{T}{n}\right)\right)^{1-\alpha} + \left(1 - \frac{T}{n}\right)\left(1 - (y-1)\frac{T}{n}\right)^{1-\alpha}. \end{aligned}$$

By a Taylor expansion of the terms $(1 \pm \cdot\,)^{1-\alpha}$ we get

$$E_{P'^n}\left[\left(1 + (y-1)\left(\zeta^{n\ell j}_k - \Delta t_k\right)\right)^{1-\alpha}\right] = 1 + \frac{(y^{1-\alpha} - y + \alpha(y-1))T + R_n}{n} \tag{5.23}$$

with $R_n = R^1_n + R^2_n + R^3_n$ for

$$\begin{aligned} R^1_n &= -\frac{T^2}{n}(1-\alpha)(1-\vartheta_1)^{-\alpha}\frac{y-1}{y^\alpha}, \quad \vartheta_1 \in \langle 0, \frac{y-1}{y}\frac{T}{n}\rangle, \\ R^2_n &= \frac{1}{2}\frac{T^2}{n}\alpha(1-\alpha)(1-\vartheta_2)^{-1-\alpha}(y-1), \quad \vartheta_2 \in \langle 0, (y-1)\frac{T}{n}\rangle, \\ R^3_n &= \frac{T^2}{n}(1-\alpha)(1-\vartheta_3)^{-\alpha}(y-1), \quad \vartheta_3 \in \langle 0, (y-1)\frac{T}{n}\rangle, \end{aligned}$$

with the general interval $\langle a, b\rangle = [\min\{a,b\}, \max\{a,b\}]$. Now the function f_α is continuous, so it attains its minimum on the compact interval $[0, K]$, and thus $f_\alpha(y) \geq f_\alpha(\bar\lambda)$ for some $0 \leq \bar\lambda \leq K$. (In fact $\bar\lambda \in \{0, K\}$ since f_α is concave.)

Let us now examine R_n. We show that $R_n \geq \frac{1}{n}\bar R$ for some $\bar R \in \mathbb{R}$. To that end we claim $nR^i_n \geq \bar R^i$ for $i \in \{1,2,3\}$. We have

1) $\frac{nR^1_n}{T^2} \geq -\frac{1}{2}(1-\alpha)(K-1)$. Indeed, if $y \leq 1$, then $\frac{nR^i_n}{T^2} = -(1-\alpha)(1-\vartheta_1)^{-\alpha}\frac{y-1}{y^\alpha} \geq 0$. If $y > 1$, then $y^\alpha > 1$, so $\frac{y-1}{y^\alpha} \leq K-1$. Furthermore in this case $\vartheta_1 \in (0, \frac{y-1}{y}\frac{T}{n})$, thus $1 - \vartheta_1 > 1 - \frac{y-1}{y}\frac{T}{n} \geq 1 - (K-1)\frac{T}{n} \geq \frac{1}{2}$ for n sufficiently large. Altogether we have the desired result.

2) $\frac{nR^2_n}{T^2} \geq -\frac{1}{2}\alpha(1-\alpha)$. In the case where $y \leq 1$ we have that $\vartheta_2 \in (\frac{(y-1)T}{n}, 0)$, so that $(1-\vartheta_2)^{-1-\alpha} < 1$, hence we have $\frac{1}{2}\alpha(1-\alpha)(1-\vartheta_1)^{-1-\alpha}(y-1) \geq -\frac{1}{2}\alpha(1-\alpha)$. In the other case where $y > 1$, it is immediate that $\frac{nR^2_n}{T^2} \geq 0$, which gives the result.

3) $\frac{nR^3_n}{T^2} \geq -(1-\alpha)$. As in 2) we see that for $y \leq 1$ we get $(1-\vartheta_3)^{-\alpha} \leq 1$, which gives $(1-\alpha)(1-\vartheta_1)^{-\alpha}(y-1) \geq -(1-\alpha)$, whereas for $y > 1$ it is again clear that $(1-\alpha)(1-\vartheta_1)^{-\alpha}(y-1) \geq 0$.

So if we set $\bar{R}_n = -\frac{T^2}{n}\left((1-\alpha)\left(\frac{1}{2}(K-1+\alpha)+1\right)\right)$, then $R_n \geq \bar{R}_n = \mathcal{O}\left(\frac{1}{n}\right)$, and (5.22) follows from (5.23) and the choice of $\bar{\lambda}$.

For n sufficiently large we now have $1 + \frac{Tf_\alpha(\bar{\lambda})+\bar{R}_n}{n} > 0$, and we get from (5.21)

$$\begin{aligned} E_{P'^n}\left[(Z_t^n)^{1-\alpha}\right] &\geq \left(1+\frac{Tf_\alpha(\bar{\lambda})+\bar{R}_n}{n}\right)^{m^2\left[\frac{nt}{T}\right]} \\ &\overset{n\to\infty}{\longrightarrow} \exp\left(tm^2 f_\alpha(\bar{\lambda})\right) \\ &\overset{\alpha\downarrow 0}{\longrightarrow} 1, \end{aligned}$$

so that $\lim_{\alpha\downarrow 0}\liminf_{n\to\infty} E_{P'^n}\left[(Z_t^n)^{1-\alpha}\right] \geq 1$, which, together with (5.20), yields the result. □

We are now in a position to state the main result of this section, namely convergence of the sequence of discrete-time models (S^n, η^n) to the continuous-time model (S, η) under P.

Theorem 5.12 *Let $L_t = t$, let (W, N, S, η) be the model (5.1)–(5.3), and let Z be given by (5.9). For the corresponding discrete-time processes $(L^n, W^n, N^n, S^n, \eta^n, Z^n)$ from (5.10), (5.11), (5.12), and (5.18), we have*

$$\mathcal{L}\left(L^n, W^n, N^n, S^n, \eta^n, Z^n \middle| P^n\right) \overset{w}{\longrightarrow} \mathcal{L}\left(L, W, N, S, \eta, Z \middle| P\right)$$

as $n \to \infty$.

Proof. We have convergence under P'^n by Proposition 5.10, and $(P^n)_{n\in\mathbb{N}} \overset{\text{loc}}{\lhd} (P'^n)_{n\in\mathbb{N}}$ by Proposition 5.11, so Theorem 1.56 yields the result. □

5.4 An Analysis of the Untradable Factors of Risk

In this section we examine properties of the processes η^n and we give a technique how to pass from P^n to P'^n. Note that under P'^n the process η^n may be viewed as a homogeneous discrete-time Markov chain with values in a finite subset of $\mathbb{Z}$. By (5.12) it is clear that all points outside of $\{1, \dots, m\}$ are absorbing, and by construction of the model it is irrelevant which value η^n takes outside of $\{1, \dots, m\}$ once it has jumped outside of $\{1, \dots, m\}$. Recall that we identify all of $\mathbb{Z} \setminus \{1, \dots, m\}$ with π and that we denote the state space of η^n by $I_m = \{\pi\} \cup \{1, \dots, m\}$, where $\eta_{t_k}^n = \pi$ means that $\eta_{t_k}^n \notin \{1, \dots, m\}$. For future reference we introduce for $\ell, j \in \{1, \dots, m\}$ the sets

$$A_k^{n\ell j} := \left\{\sum_{i=1}^m (i-\ell)\ \zeta_k^{n\ell i} = j - \ell\right\} = \left\{\ell + \sum_{i=1}^m (i-\ell)\ \zeta_k^{n\ell i} = j\right\}$$

and

$$A_k^{n\ell\pi} := \left\{\ell + \sum_{i=1}^m (i-\ell)\ \zeta_k^{n\ell i} \notin \{1, \dots, m\}\right\}.$$

For notational convenience we also define the sets $A_k^{n\pi j}$, where $P'^n\big[A_k^{n\pi j}\big] = \delta_{\pi j}$. This corresponds to the fact that π is absorbing, and $P^n \ll P'^n$ yields that $P^n\big[A_k^{n\pi j}\big] = \delta_{\pi j}$.

For fixed n and $\ell, j \in I_m$ we denote the transition probabilities of η^n under P'^n by $p'^{n\ell j}$, and we have for $\ell \in \{1, \dots, m\}$, $j \in I_m$

$$p'^{n\ell j} = P'^n\left[\eta^n_{t_k} = j \;\middle|\; \eta^n_{t_{k-1}} = \ell\right] = P'^n\left[A_k^{n\ell j}\right]. \tag{5.24}$$

It is immediate that $p'^{n\ell\pi} = 1 - \sum_{j=1}^m p'^{n\ell j}$, and we have the following properties of $p'^{n\ell j}$.

Lemma 5.13 *Let $p'^{n\ell j}$ be given as above. Then there exists a constant c such that for $\ell, j \in I_m$ and n sufficiently large*

a) $p'^{n\ell\ell} \geq 1 - \frac{c}{n}$,

b) $p'^{n\ell j} \leq \frac{c}{n}$ *for* $j \neq \ell$,

c) $p'^{n\ell\pi} \leq \frac{c}{n^2}$.

Proof. a) and b) are special cases of Lemma 5.14 below, whereas the proof of c) is a simple but lengthy computation. It is therefore relegated to the Appendix; see Section C. □

Notice that Lemma 5.13 c) yields that the P'^n-probability for η^n to leave the set $\{1, \dots, m\}$ vanishes as n tends to infinity. Since $P^n \ll P'^n$ for all $n \in \mathbb{N}$, we conclude that the corresponding P^n-probability tends to 0 as well.

The above properties of η^n, in particular the time-homogeneity of η^n under P'^n, result from the independence of the $\zeta_k^{n\ell i}$ under P'^n. Obviously under P^n this is not the case any more. However we have the following result concerning the transition probabilities. Recall that for $j \in I_m$ we have sets $A_k^{n\pi j}$, where $P^n\big[A_k^{n\pi j}\big] = \delta_{\pi j}$. Therefore we have for $\ell, j \in I_m$

$$\mathbb{1}_{\{\eta^n_{t_k} = j, \eta^n_{t_{k-1}} = \ell\}} = \mathbb{1}_{A_k^{n\ell j} \cap \{\eta^n_{t_{k-1}} = \ell\}} \quad P^n\text{-a.s.}$$

Recall furthermore the density process Z^n from (5.19). By the Bayes formula for conditional expectation we get for $j \in I_m$

$$\begin{aligned}
&P^n\left[\eta^n_{t_k} = j \,\middle|\, \mathcal{F}^n_{t_{k-1}}\right] \\
&= \sum_{\ell \in I_m} E_{P^n}\left[\mathbb{1}_{\{\eta^n_{t_k} = j, \eta^n_{t_{k-1}} = \ell\}} \,\middle|\, \mathcal{F}^n_{t_{k-1}}\right] \\
&= \sum_{\ell \in I_m} E_{P'^n}\left[\mathbb{1}_{A_k^{n\ell j} \cap \{\eta^n_{t_{k-1}} = \ell\}} \prod_{j_1, j_2 = 1}^m \left(1 + \left(\lambda^{j_1 j_2}(t_k, S^n_{t_{k-1}}) - 1\right)\left(\zeta_k^{n j_1 j_2} - \frac{T}{n}\right)\right) \,\middle|\, \mathcal{F}^n_{t_{k-1}}\right] \\
&= \sum_{\ell \in I_m} E_{P'^n}\left[\mathbb{1}_{A_k^{n\ell j}} \prod_{j_1, j_2 = 1}^m \left(1 + \left(\lambda^{j_1 j_2}(t_k, x) - 1\right)\left(\zeta_k^{n j_1 j_2} - \frac{T}{n}\right)\right)\right]\Bigg|_{x = S^n_{t_{k-1}}} \mathbb{1}_{\{\eta^n_{t_{k-1}} = \ell\}} \\
&=: \sum_{\ell \in I_m} p_k^{n\ell j}(S^n_{t_{k-1}})\ \mathbb{1}_{\{\eta^n_{t_{k-1}} = \ell\}},
\end{aligned}$$

Notice that $\mathbb{1}_{A_k^{n\ell j}}$ is a measurable function of $\zeta_k^{n\ell i}$, $i \in \{1,\dots,m\}$, only, so due to the fact that the $\zeta_k^{nj_1j_2}$ are independent under P'^n and $E_{P'^n}\left[1 + \left(\lambda^{j_1j_2}(t_k,x) - 1\right)\left(\zeta_k^{nj_1j_2} - \frac{T}{n}\right)\right] = 1$, we actually have for $\ell, j \in \{1,\dots,m\}$ and $x \in \mathbb{R}^d$

$$p_k^{n\ell j}(x) = E_{P'^n}\left[\mathbb{1}_{A_k^{n\ell j}} \prod_{j_2=1}^{m}\left(1 + \left(\lambda^{\ell j_2}(t_k,x) - 1\right)\left(\zeta_k^{n\ell j_2} - \frac{T}{n}\right)\right)\right]. \tag{5.25}$$

This uses that the factors with $j_1 \neq \ell$ are independent of $A_k^{n\ell j}$ and have expectation 1. For $j \in I_m$, $\ell \in \{1,\dots,m\}$, the above definition of $p_k^{n\ell j}$ and $P^n\left[A_k^{n\pi j}\right] = \delta_{\pi j}$ yields

$$\begin{cases} p_k^{n\pi j}(x) &=& \delta_{\pi j}, \\ p_k^{n\ell\pi}(x) &=& 1 - \sum_{i=1}^m p_k^{n\ell i}(x). \end{cases} \tag{5.26}$$

The transition probabilities of η^n as defined above give us a convenient tool to convert conditional expectations with respect to P^n into conditional expectations with respect to P'^n. Let us first state some useful properties of $p_k^{n\ell j}(x)$.

Lemma 5.14 *Let $n \in \mathbb{N}$, $k \in \{1,\dots,n\}$ and $p_k^{n\ell j}(x)$ be defined by (5.25)–(5.26). Then there exists a constant c, independent of n, such that for all $\ell, j \in I_m$ the following hold:*

a) $p_k^{n\ell\ell}(x) \geq 1 - \frac{c}{n}$ for all $x \in \mathbb{R}^d$,

b) $p_k^{n\ell j}(x) \leq \frac{c}{n}$ for all $x \in \mathbb{R}^d$ and $j \neq \ell$,

c) $p_k^{n\ell j}$ is C^1, and $|\nabla p_k^{n\ell j}(x)| \leq \frac{c}{n}$.

Proof. a) For $\ell = \pi$ we have $p_k^{n\pi\pi}(x) = 1$ by definition. For $\ell \in \{1,\dots,m\}$ notice that $\bigcap_{i=1}^m \{\zeta_k^{n\ell i} = 0\} \subseteq A_k^{n\ell\ell}$, and that by assumption $\lambda^{\ell j}(t,x) \leq c_0$, uniformly in t and x. So we have by the independence of $\zeta_k^{n\ell j}$ under P'^n that

$$\begin{aligned} p_k^{n\ell\ell}(x) &\geq E_{P'^n}\left[\prod_{i=1}^m \mathbb{1}_{\{\zeta_k^{n\ell i}=0\}} \prod_{j=1}^m \left(1 + \left(\lambda^{\ell j}(t_k,x) - 1\right)\left(\zeta_k^{n\ell j} - \frac{T}{n}\right)\right)\right] \\ &= E_{P'^n}\left[\prod_{i=1}^m \mathbb{1}_{\{\zeta_k^{n\ell i}=0\}} \left(1 - \left(\lambda^{\ell i}(t_k,x) - 1\right)\frac{T}{n}\right)\right] \\ &= \prod_{i=1}^m P'^n\left[\zeta_k^{n\ell i} = 0\right]\left(1 - \left(\lambda^{\ell i}(t_k,x) - 1\right)\frac{T}{n}\right) \\ &\geq \left(1 - \frac{T}{n}\right)^m \left(1 - (c_0 - 1)\frac{T}{n}\right)^m \\ &\geq 1 - \frac{c'}{n} \end{aligned}$$

for c' chosen appropriately.

b) This follows from a) and the fact that $\sum_{i\in I_m} p_k^{n\ell i}(x) = 1$, because for $j \neq \ell$ we have

$$p_k^{n\ell j}(x) \le \sum_{\substack{i\in I_m\\ i\neq \ell}}^{m} p_k^{n\ell i}(x) = 1 - p_k^{n\ell\ell}(x) \le \frac{c'}{n}.$$

c) Differentiability of $p_k^{n\ell j}$ follows from Lemma C.7 in the Appendix, and it remains to prove the boundedness of the gradient of $p_k^{n\ell j}$. Note that $p_k^{n\pi j}$ is constant for all $j \in I_m$, and boundedness of the gradient of $p_k^{n\ell\pi} = 1-\sum_{j=1}^{m} p_k^{n\ell j}(x)$ follows from the boundedness of $|\nabla p_k^{n\ell j}|$ for $j \in \{1,\dots,m\}$. So we fix $n \in \mathbb{N}$, $k \in \{1,\dots,n\}$, $\ell, j \in \{1,\dots,m\}$ and $r \in \{1,\dots,d\}$, and we have by the product rule

$$\begin{aligned}
&\left|\frac{\partial}{\partial x^r} p_k^{n\ell j}(x)\right| \\
&= \left|\sum_{i=1}^{m} E_{P'^n}\left[\mathbb{1}_{A_k^{n\ell j}} \frac{\partial \lambda^{\ell i}}{\partial x^r}(t_k,x)\left(\zeta_k^{n\ell i} - \frac{T}{n}\right)\prod_{\substack{s=1\\ s\neq i}}^{m}\left(1 + \left(\lambda^{\ell s}(t_k,x) - 1\right)\left(\zeta_k^{n\ell s} - \frac{T}{n}\right)\right)\right]\right| \\
&\le \sum_{i=1}^{m}\left|\frac{\partial \lambda^{\ell i}}{\partial x^r}(t_k,x)\right| E_{P'^n}\left[\left|\zeta_k^{n\ell i} - \frac{T}{n}\right|\right]\prod_{\substack{s=1\\ s\neq i}}^{m} E_{P'^n}\left[\left|1 + \left(\lambda^{\ell s}(t_k,x) - 1\right)\left(\zeta_k^{n\ell s} - \frac{T}{n}\right)\right|\right].
\end{aligned}$$

Now from the distribution of $\zeta_k^{n\ell j}$ under P'^n it is immediate to see that there exist constants c_1 and c_2 such that

$$E_{P'^n}\left[\left|\zeta_k^{n\ell i} - \frac{T}{n}\right|\right] = 2\left(1 - \frac{T}{n}\right)\frac{T}{n} \le \frac{c_1}{n}$$

and

$$E_{P'^n}\left[\left|1 + \left(\lambda^{\ell s}(t_k,x) - 1\right)\left(\zeta_k^{n\ell s} - \frac{T}{n}\right)\right|\right] \le 1 + \left|\lambda^{\ell s}(t_k,x) - 1\right|\left(1 - \frac{T}{n}\right)\frac{T}{n} \le c_2,$$

so that the boundedness of $|\lambda^{\ell j}|$ and $|\nabla\lambda^{\ell j}|$ by the constant c_0 implies $\left|\frac{\partial}{\partial x^r} p_k^{n\ell j}(x)\right| \le \frac{c''}{n}$ for $c'' = m\ c_0\ c_1\ c_2^{m-1}$. Finally from $p_k^{n\ell\pi} = 1 - \sum_{j=1}^{m} p_k^{n\ell j}$ we have $|\nabla p_k^{n\ell\pi}| \le m\frac{c''}{n}$, so $c = c' \vee mc''$ is good enough. □

Intuitively, the change of measure from P'^n to P^n involves only the "intensities" of N^n, the processes driving η^n. Since η^n influences the evolution of S^n as well, we cannot expect that the change of measure should have no influence on S^n. However we have the following useful result concerning the change of conditional expectations under the change of measure from P'^n to P^n.

Lemma 5.15 *Let* $f\colon (\mathbb{R}^d)^2 \times (I_m)^2 \to \mathbb{R}$ *be measurable. Then*

$$\begin{aligned}
&E_{P^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}, \eta^n_{t_k}, \eta^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] \\
&= \sum_{\ell,j\in I_m} E_{P'^n}\big[f(x+\Sigma(t_k,x,\ell)\xi^n_k + \Gamma(t_k,x,\ell)\Delta t_k, x, j, \ell)\big]\big|_{x=S^n_{t_{k-1}}} \; p^{n\ell j}_k(S^n_{t_{k-1}}) \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}} \\
&= \sum_{\ell,j\in I_m} E_{P'^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}, j, \eta^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] \; p^{n\ell j}_k(S^n_{t_{k-1}}) \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}}.
\end{aligned}$$

Proof. Note that for fixed k the random variables ξ^n_k and $(\zeta^{n\ell j}_k)_{\ell,j\in\{1,\dots,m\}}$ are independent and independent of $\mathcal{F}^n_{t_{k-1}}$ under P'^n. Recall that $\mathbb{1}_{\{\eta^n_{t_k}=j,\eta^n_{t_{k-1}}=\ell\}} = \mathbb{1}_{A^{n\ell j}_k \cap \{\eta^n_{t_{k-1}}=\ell\}}$ P^n-a.s. for $\ell, j \in I_m$. Then we get by the Bayes formula for conditional expectations and the fact that $P^n\big[A^{n\pi j}_k\big] = \delta_{\pi j}$ for $j \in I_m$

$$\begin{aligned}
&E_{P^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}, \eta^n_{t_k}, \eta^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] \\
&= \sum_{\ell,j\in I_m} E_{P^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}, j, \ell) \; \mathbb{1}_{\{\eta^n_{t_k}=j,\eta^n_{t_{k-1}}=\ell\}} \middle| \mathcal{F}^n_{t_{k-1}} \right] \\
&= \sum_{\ell,j\in I_m} E_{P'^n}\Bigg[f(S^n_{t_k}, S^n_{t_{k-1}}, j, \ell) \; \mathbb{1}_{A^{n\ell j}_k} \times \\
&\qquad \times \prod_{j_1,j_2=1}^{m} \left(1 + \left(\lambda^{j_1 j_2}(t_k, S^n_{t_{k-1}}) - 1\right)\left(\zeta^{n j_1 j_2}_k - \frac{T}{n}\right)\right) \Bigg| \mathcal{F}^n_{t_{k-1}} \Bigg] \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}} \\
&= \sum_{\ell,j\in I_m} E_{P'^n}\big[f(x+\Sigma(t_k,x,\ell)\xi^n_k + \Gamma(t_k,x,\ell)\Delta t_k, x, j, \ell)\big]\big|_{x=S^n_{t_{k-1}}} \; p^{n\ell j}_k(S^n_{t_{k-1}}) \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}} \\
&= \sum_{\ell,j\in I_m} E_{P'^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}, j, \eta^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] p^{n\ell j}_k(S^n_{t_{k-1}}) \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}},
\end{aligned}$$

which shows the claimed equalities. □

Remark 5.16 a) Note that Lemma 5.14 and Lemma 5.15 include the corresponding results for P'^n: If we set $\lambda \equiv 1$ in (5.26), then $P^n = P'^n$ and $p^{n\ell j}_k(x) = p'^{n\ell j}$.

b) Moreover Lemma 5.15 yields that the conditional distributions under P^n and P'^n of $(S^n_{t_k}, S^n_{t_{k-1}})$ given $\mathcal{F}^n_{t_{k-1}}$ coincide in the sense that for f measurable

$$\begin{aligned}
E_{P^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] &= \sum_{\ell,j\in I_m} E_{P'^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] p^{n\ell j}_k(S^n_{t_{k-1}}) \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}} \\
&= E_{P'^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right] \sum_{\ell,j\in I_m} p^{n\ell j}_k(S^n_{t_{k-1}}) \; \mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}} \\
&= E_{P'^n}\left[f(S^n_{t_k}, S^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}} \right],
\end{aligned}$$

since for fixed ℓ we have $\sum_{j\in I_m} p^{n\ell j}_k(S^n_{t_{k-1}}) = 1$ by (5.26). This is also intuitively clear since $S^n_{t_k}$ is determined by $S^n_{t_{k-1}}$, $\eta^n_{t_{k-1}}$ and ξ^n_k. The first two are $\mathcal{F}^n_{t_{k-1}}$-measurable, and ξ^n_k is independent of $\mathcal{F}^n_{t_{k-1}}$ with the same distribution under both P^n and P'^n because the change of measure only affects N^n, the process driving η^n. $\diamond$

Chapter 6

Convergence of Price Processes

In this chapter we investigate convergence results in the setting of *pure pricing* for the model (S, η) from Chapter 5. Note that in general the model is incomplete due to the non-tradable risks η so that there are many different martingale measures for S. Pure pricing then means that we start with an a priori fixed martingale measure for S, which we use for pricing. In our model P itself is a (local) martingale measure for S if $\Gamma \equiv 0$ in (5.1). Thus we assume that (S, η) is given by

$$dS_t = \Sigma(t, S_t, \eta_{t-})\, dW_t, \quad S_0 \in \mathbb{R}^d, \tag{6.1}$$

$$d\eta_t = \sum_{\ell,j=1}^{m} (j-\ell)\mathbb{1}_{\{\eta_{t-}=\ell\}}\, dN_t^{\ell j}, \quad \eta_0 \in \{1, \dots, m\}, \tag{6.2}$$

where W is an r-dimensional standard Brownian motion and N is a multivariate point process satisfying (5.3) under P. The same structure also appears from other approaches to pricing and hedging. We could for instance start with (S, η) given by (5.1)–(5.3) under some measure P; then one can show that under the minimal martingale measure $\hat{P}$ from Föllmer and Schweizer (1990) and Schweizer (1995), (S, η) has the dynamics (6.1)–(6.2). Alternatively, we could begin with $\lambda^{\ell j} \equiv 1$ for $\ell, j \in \{1, \dots, m\}$ and use a further development of Section 4.4 to deduce that under the entropy-minimizing martingale measure, (S, η) has again the dynamics (6.1)–(6.2). Hence the setting (6.1)–(6.2) is a natural starting point to analyze the behaviour of a pricing mechanism.

Our interest in this chapter lies in approximating a price process V, which is given as the solution of a backward stochastic differential equation, by a sequence of discrete-time processes V^n and in the sense of convergence in distribution, i.e. $\mathcal{L}(V^n|P^n) \xrightarrow{w} \mathcal{L}(V|P)$. The usual procedure to prove convergence in distribution is to show first that the sequence of approximating distributions is tight, and then to identify every cluster point of the sequence with the desired limit. But in our case the situation is more complicated because the limiting process V as well as the approximating processes V^n are not given explicitly, but defined implicitly as solutions of backward stochastic differential equations. Since these involve conditional expectations, we have to take filtrations into account as well.

In Section 6.1 we present the models for price processes in continuous and discrete time and specify the conditions under which we later show $\mathcal{L}(V^n|P^n) \xrightarrow{w} \mathcal{L}(V|P)$. In Section 6.2 we lay out a computation scheme that gives us a representation $V_t^n = v^n(t, S_t^n, \eta_t^n)$, where v^n satisfies certain regularity conditions. These allow us in Section 6.3 to show tightness of the sequence $\mathcal{L}(V^n|P'^n)$ with the help of Theorem 1.45.

In Section 6.4 we identify every cluster point of the approximating sequence with a certain process V. This V crucially depends on the chosen filtration in the continuous-time model, whereas convergence in distribution has nothing to do with filtrations. This difficulty can be partly circumvented by using the concept of *convergence of filtrations* as presented by Coquet, Mémin and Słominski (2001). Because convergence of filtrations involves processes which converge in probability and filtrations which all live on the same probability space, it will be useful to apply the Skorokhod embedding theorem which provides exactly such a situation. The remaining problem is that there are only few situations where the convergence of a given sequence of filtrations can be shown with reasonable effort. One such case occurs when the filtrations are generated by processes with independent increments. This is why we show tightness and convergence first under P'^n, P' and then apply Theorem 1.56 in order to get convergence under P^n and P. Note, however, that V^n and V are *always* defined via conditional expectations with respect to P^n and P.

Throughout this chapter we assume that in the continuous-time case we are given a probability space $(\Omega, \mathcal{F}, P')$ with an r-dimensional P'-standard Brownian motion W and a multivariate P'-standard Poisson process N. We define the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]} = \mathbb{F}^{(W,N)}(P')$ to be the P'-augmentation of the filtration generated by W and N, and we furthermore assume that $dP = Z_T \, dP'$ as in (5.8). Concerning the approximating models we assume that for each $n \in \mathbb{N}$ the probability space $(\Omega^n, \mathcal{F}^n, P'^n)$ is endowed with $\mathbb{F}^n = (\mathcal{F}_t^n)_{t\in[0,T]} = \mathbb{F}^{(W^n,N^n)}(P'^n)$.

6.1 The Pure Pricing Approach in Continuous and Discrete Time

In the approach of pure pricing one *defines* the price of a contingent claim as the conditional expectation under the pricing measure P of all future payments. For the continuous-time model (S, η) given by (5.1)–(5.3) with $\Gamma \equiv 0$ this is investigated by Becherer and Schweizer (2003) who consider payment structures of the form

$$B = h(S_T, \eta_T) + \int_0^T \delta(s, S_s, \eta_{s-}, V_{s-}) \, ds + \int_0^T \sum_{\ell,j=1}^{m} f^{\ell j}(s, S_s, V_{s-}) \, dN_s^{\ell j}, \tag{6.3}$$

where V is the price process of B, and where $h\colon \mathbb{R}^d \times I_m \to \mathbb{R}$, $\delta\colon [0,T] \times \mathbb{R}^d \times I_m \times \mathbb{R} \to \mathbb{R}$, and $f^{ij}\colon [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ are bounded functions, and $h(\cdot, \ell)$, $\delta(\cdot, \cdot, \ell, \cdot)$ and f^{ij} are C^1 for all i, j and ℓ and locally Lipschitz in the argument v, uniformly in (t, x). Note that in the continuous-time case η never leaves $\{1,\dots,m\}$, so the dependence of h and δ on the cemetery point π in $I_m = \{\pi\} \cup \{1,\dots,m\}$ may be neglected.

If we interpret η as the evolution of a rating for some asset S, then B can be interpreted as a composition of several payments. h is a payment made at expiration time T, which depends on the final value of both stock and rating; δ models some continuously made payments, which depend on the current values of S, η *and* V, whereas the amount $f^{\ell j}$ is made every time the rating jumps from state ℓ to state j. The dependence of B on its own current price occurs, e.g., with fractional recovery of defaultable bonds.

Note that due to the dependence of B on its own price process, B is not yet well-defined. To overcome this problem it is convenient to first define the price process V as the conditional expectation of all future payments under the pricing measure. This leads to the backward stochastic differential equation

$$V_t = E_P\left[h(S_T,\eta_T) + \int_t^T \delta(u,S_u,\eta_{u-},V_{u-})\,du + \int_t^T \sum_{\ell,j=1}^m f^{\ell j}(u,S_u,V_{u-})\,dN_u^{\ell j} \,\middle|\, \mathcal{F}_t \right] \tag{6.4}$$

for $t \in [0,T]$, where $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is the P'-augmentation of the filtration generated by (W,N). If (6.4) admits a unique solution process V, then B can be defined as in (6.3).

Remark 6.1 Becherer and Schweizer (2003) show that under the above conditions on h, δ and f the recursion formula (6.4) admits a unique solution V in the class of bounded semimartingales, and that $V_t = v(t,S_t,\eta_t)$, where v solves for each $\ell \in \{1,\dots,m\}$ the reaction-diffusion equation

$$\begin{aligned} 0 \;=\; & \frac{\partial v}{\partial t}(t,x,\ell) + \frac{1}{2}\sum_{i,k=1}^d a^{ik}(t,x,\ell)\frac{\partial^2 v}{\partial x^i \partial x^k}(t,x,\ell) + \delta(t,x,\ell,v(t,x,\ell)) \\ & + \sum_{\substack{j=1\\ j\neq \ell}} \lambda^{\ell j}(t,x)\Big(v(t,x,j) - v(t,x,\ell) + f^{\ell j}(t,x,v(t,x,\ell))\Big) \end{aligned} \tag{6.5}$$

for $(t,x) \in [0,T) \times (0,\infty)^d$ with the boundary condition $v(T,x,\ell) = h(x,\ell)$. In (6.5) we use the notation $a = (a^{ik})_{i,k=1,\dots,d} := \Sigma\Sigma^{\mathrm{tr}}$. ◇

In a discrete-time setting the analogous approach is as follows. For $n \in \mathbb{N}$ and $k \in \{0,\dots,n\}$ let $t_k := k\frac{T}{n}$ and $\Delta t_k := t_k - t_{k-1} = \frac{T}{n}$. Let the discrete-time processes S^n and η^n be defined as in Section 5.2 for $\Gamma \equiv 0$, i.e. S^n and η^n are piecewise constant processes with $S_0^n = S_0 \in \mathbb{R}^d, \eta_0^n = \eta_0 \in \{1,\dots,m\}$ and

$$\Delta S_{t_k}^n \;=\; \Sigma(t_k, S_{t_{k-1}}^n, \eta_{t_{k-1}}^n)\,\Delta W_{t_k}^n, \tag{6.6}$$

$$\Delta \eta_{t_k}^n \;=\; \sum_{\ell,j=1}^m (j-\ell)\mathbb{1}_{\{\eta_{t_{k-1}}^n = \ell\}}\,\Delta N_{t_k}^{n\ell j}, \tag{6.7}$$

where W^{ni} and $N^{n\ell j}$ for $i \in \{1,\dots,r\}$ and $\ell, j \in \{1,\dots,m\}$ are the binomial processes defined in (5.10). Recall that by Lemma 5.8 the processes $N^{n\ell j}$ have P^n-"intensities" $\bar\lambda^{n\ell j}(t,x) = \lambda^{\ell j}(t,x) + \frac{T}{n}(1-\lambda^{\ell j}(t,x))$ in the sense that

$$E_{P^n}\left[\Delta N^{n\ell j}_{t_k}\middle|\mathcal{F}^n_{t_{k-1}}\right] = \bar\lambda^{n\ell j}(t_k, S^n_{t_{k-1}})\ \Delta t_k.$$

In analogy to the continuous-time case the filtration $\mathbb{F}^n = (\mathcal{F}^n_t)_{t\in[0,T]} = \mathbb{F}^{(W^n,N^n)}(P^n)$ is generated by W^n and N^n. We then define the price process V^n as a solution of

$$\begin{aligned} V^n_{t_k} &= E_{P^n}\Bigg[h(S^n_T,\eta^n_T) + \sum_{i=k+1}^{n} \delta(t_i, S^n_{t_{i-1}}, \eta^n_{t_{i-1}}, V^n_{t_{i-1}})\ \Delta t_i \\ &\qquad + \sum_{i=k+1}^{n}\sum_{\ell,j=1}^{m} f^{\ell j}(t_i, S^n_{t_{i-1}}, V^n_{t_{i-1}})\ \Delta N^{n\ell j}_{t_i}\Bigg|\mathcal{F}^n_{t_k}\Bigg] \end{aligned} \tag{6.8}$$

for $k \in \{0,\dots,n\}$ and $V^n_t = V^n_{t_k}$ for $t_k \le t < t_{k+1}$. At last the payment structure under consideration in the discrete-time is given by

$$B^n = h(S^n_T,\eta^n_T) + \sum_{i=1}^{n} \delta(t_i, S^n_{t_{i-1}}, \eta^n_{t_{i-1}}, V^n_{t_{i-1}})\ \Delta t_i + \sum_{i=1}^{n}\sum_{\ell,j=1}^{m} f^{\ell j}(t_i, S^n_{t_{i-1}}, V^n_{t_{i-1}})\ \Delta N^{n\ell j}_{t_i}. \tag{6.9}$$

Recall that η^n may jump out of the set $\{1,\dots,m\}$ but remains constant thereafter, and that we identify all points in the range of η^n outside of $\{1,\dots,m\}$ with the cemetery point π.

Below we will show that for all $n \in \mathbb{N}$ there is at most one solution of (6.8) under the above conditions on h, δ and $f^{\ell j}$. To prove existence of a solution for $n \in \mathbb{N}$ and convergence of the sequence of solutions to V, which will be discussed in Sections 6.2–6.4, we need some more assumptions on h, δ, $f^{\ell j}$ and $\lambda^{\ell j}$, or more precisely on their derivatives. To ease notation we first define for the remaining part of this chapter for $t \in [0,T]$, $x \in \mathbb{R}^d$, $\ell \in I_m$, and $v \in \mathbb{R}$

$$\left\{\begin{aligned} \delta^\ell &= \delta(t,x,\ell,v), \\ \gamma^n(t,x,\ell,v) &= \delta^\ell(t,x,v) + \textstyle\sum_{i,j=1}^m f^{ij}(t,x,v)\ \bar\lambda^{nij}(t,x), \\ \gamma^{n\ell}(t,x,v) &= \gamma^n(t,x,\ell,v), \\ h^\ell(x) &= h(x,\ell). \end{aligned}\right. \tag{6.10}$$

Furthermore we define for $\ell \in I_m$ the function $\Sigma^\ell\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times r}$ by $\Sigma^\ell(t,x) = \Sigma(t,x,\ell)$.

Now we assume in addition that for all $\ell \in I_m$ and $i,j \in \{1,\dots,m\}$

$$h^\ell, \delta^\ell, f^{ij} \text{ and } \lambda^{ij} \text{ are bounded and admit a bounded gradient, and } \lambda^{ij}(t,x) > 0 \tag{6.11}$$

for all $t \in [0,T]$ and $x \in \mathbb{R}^d$.

The condition on λ^{ij} implies that the density processes Z^n and Z from (5.19) and (5.9) are strictly positive P'^n- and P'-martingales, respectively, and thus $P^n \sim P'^n$ and $P \sim P'$. We need equivalence of P and P' later on when we identify V as the only possible cluster point of a subsequence of V^n. Note that by the definitions of $\bar{\lambda}^{nij}$ and $\gamma^{n\ell}$ (6.11) implies that $\gamma^{n\ell}$ is in C^1 and uniformly bounded with uniformly bounded gradient and thus in particular globally Lipschitz, uniformly in n. To simplify notation we choose c_0 large enough such that $|\nabla h^\ell|$, $|\nabla \lambda^{ij}|$ as well as $|\nabla \gamma^{n\ell}|$ and $|\nabla \gamma^{n\ell}|_{\max}$ are bounded by c_0, uniformly in $n \in \mathbb{N}$. Here $|\cdot|_{\max}$ denotes the maximum norm on $\mathbb{R}^{d'}$ for arbitrary $d' \in \mathbb{N}$. Furthermore we assume that Σ has bounded derivatives as well in the sense that $|\frac{\partial \Sigma^\ell}{\partial x^k}| \leq \bar{\Sigma}$, uniformly in t and x, for some constant $\bar{\Sigma}$.

We start with the following preparatory lemma which reduces the question about existence and uniqueness of a solution process of (6.8) to a local question.

Lemma 6.2 *Suppose that the recursion formula*

$$\left\{ \begin{array}{lcl} V_T^n &=& h(S_T^n, \eta_T^n), \\ V_{t_k}^n &=& E_{P^n}\left[V_{t_{k+1}}^n \middle| \mathcal{F}_{t_k}^n \right] + \gamma^n\left(t_{k+1}, S_{t_k}^n, \eta_{t_k}^n, V_{t_k}^n\right)\ \Delta t_{k+1} \end{array} \right. \tag{6.12}$$

for $k \in \{0, \ldots, n-1\}$ and $V_t^n = V_{t_k}^n$ for $t_k \leq t < t_{k+1}$, admits a unique solution process V^n. Then V^n satisfies (6.8). Conversely if V^n satisfies (6.8), then V^n is given by (6.12).

Proof. To ease notation we write $\delta_{t_k}^n := \delta(t_{k+1}, S_{t_k}^n, \eta_{t_k}^n, V_{t_k}^n)$ and $f_{t_k}^{n\ell j} := f^{\ell j}(t_{k+1}, S_{t_k}^n, V_{t_k}^n)$. Recall that $E_{P^n}\left[\Delta N_{t_{k+1}}^{n\ell j} | \mathcal{F}_{t_k}^n\right] = \bar{\lambda}^{n\ell j}(t_{k+1}, S_{t_k}^n)\Delta t_{k+1}$ for $\ell, j \in \{1, \ldots, m\}$, so (6.8) yields

$$\begin{aligned} V_{t_k}^n &= E_{P^n}\left[h(S_T^n, \eta_T^n) + \sum_{i=k+2}^{n} \delta_{t_{i-1}}^n \Delta t_i + \sum_{i=k+2}^{n} \sum_{\ell,j=1}^{m} f_{t_{i-1}}^{n\ell j}\ \Delta N_{t_i}^{n\ell j} \middle| \mathcal{F}_{t_k}^n \right] \\ &\quad + \delta_{t_{k+1}}^n \Delta t_k + \sum_{\ell,j=1}^{m} f_{t_k}^{n\ell j}\ E\left[\Delta N_{t_{k+1}}^{n\ell j} \middle| \mathcal{F}_{t_k}^n \right] \\ &= E_{P^n}[V_{t_{k+1}}^n | \mathcal{F}_{t_k}^n] \\ &\quad + \left(\delta(t_{k+1}, S_{t_k}^n, \eta_{t_k}^n, V_{t_k}^n)\ + \sum_{\ell,j=1}^{m} f^{\ell j}(t_{k+1}, S_{t_k}^n, V_{t_k}^n)\ \bar{\lambda}^{n\ell j}(t_{k+1}, S_{t_k}^n) \right) \Delta t_k, \end{aligned}$$

by conditioning on $\mathcal{F}_{t_{k+1}}^n$. Hence (6.8) implies (6.12). For the converse implication we use a backward induction argument. Suppose there exists a process V^n which satisfies (6.12) for all $k \in \{0, \ldots, n-1\}$. Then $V_{t_k}^n$ obviously satisfies (6.8) for $k = n$. Now suppose $V_{t_k}^n$ satisfies

(6.8) for some $k \le n$, then by (6.12) we have

$$\begin{aligned} V^n_{t_{k-1}} &= E_{P^n}\left[V^n_{t_k}\middle|\mathcal{F}^n_{t_{k-1}}\right] + \gamma^n\left(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, V^n_{t_{k-1}}\right)\Delta t_k \\ &= E_{P^n}\left[h(S^n_T,\eta^n_T) + \sum_{i=k+1}^{n}\delta^n_{t_{i-1}}\Delta t_i + \sum_{i=k+1}^{n}\sum_{\ell,j=1}^{m} f^{n\ell j}_{t_{i-1}}\ \Delta N^{n\ell j}_{t_i}\middle|\mathcal{F}^n_{t_{k-1}}\right] \\ &\quad +\gamma^n\left(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, V^n_{t_{k-1}}\right)\Delta t_k \\ &= E_{P^n}\left[h(S^n_T,\eta^n_T) + \sum_{i=k}^{n}\delta^n_{t_{i-1}}\Delta t_i + \sum_{i=k}^{n}\sum_{\ell,j=1}^{m} f^{n\ell j}_{t_{i-1}}\ \Delta N^{n\ell j}_{t_i}\middle|\mathcal{F}^n_{t_{k-1}}\right], \end{aligned}$$

since $\gamma^n\left(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, V^n_{t_{k-1}}\right)$ is $\mathcal{F}^n_{t_{k-1}}$-measurable, and

$$\gamma^n\left(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, V^n_{t_{k-1}}\right)\Delta t_k = E_{P^n}\left[\delta^n_{t_{k-1}}\Delta t_k + \sum_{\ell,j=1}^{m} f^{n\ell j}_{t_{k-1}}\ \Delta N^{n\ell j}_{t_k}\middle|\mathcal{F}^n_{t_{k-1}}\right].$$

This finishes the proof. □

Proposition 6.3 *Let V^n and $\bar{V}^n$ be two solutions of the backward stochastic differential equation (6.8). Then $V^n_{t_k} = \bar{V}^n_{t_k}$ P-a.s. for all $k \in \{0,\dots,n\}$ and n sufficiently large.*

Proof. Recall that γ^n is Lipschitz in v, uniformly in t, x and n, so that there exists a constant $L_\gamma < \infty$ such that $|\gamma^n(t,x,\ell,v) - \gamma^n(t,x,\ell,\bar{v})| \le L_\gamma|v-\bar{v}|$ for all t, x, ℓ and n. Choose n large enough so that $\frac{T}{n}L_\gamma < 1$.

We now use induction over k. Since V^n and $\bar{V}^n$ both solve (6.8), they also solve (6.12). In particular, we have $V^n_T = h(S^n_T, \eta^n_T) = \bar{V}^n_T$ and the assertion is shown for $k = n$. So suppose that $V^n_{t_k} = \bar{V}^n_{t_k}$ for some $k \le n$, which obviously yields that $E_{P^n}\left[V^n_{t_k}\middle|\mathcal{F}^n_{t_{k-1}}\right] = E_{P^n}\left[\bar{V}^n_{t_k}\middle|\mathcal{F}^n_{t_{k-1}}\right]$. Then we have by (6.12)

$$\begin{aligned} \left|V^n_{t_{k-1}} - \bar{V}^n_{t_{k-1}}\right| &= \left|\gamma^n(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, V^n_{t_{k-1}}) - \gamma^n(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, \bar{V}^n_{t_{k-1}})\right| \Delta t_k \\ &\le \frac{T}{n}L_\gamma\left|V^n_{t_{k-1}} - \bar{V}^n_{t_{k-1}}\right|. \end{aligned}$$

So if $V^n_{t_{k-1}} \ne \bar{V}^n_{t_{k-1}}$, we get a contradiction. □

So far we have not shown existence of a solution V^n of the discrete-time backward stochastic differential equation (6.8) or, equivalently, the recursion formula (6.12). In the next section we develop a computation scheme to solve (6.8).

6.2 A Computation Scheme for the Price Process in Discrete Time

We now come to a central point on our track to show tightness and convergence of V^n. Note, however, that we do not even know yet if the discrete-time backward stochastic differential equation (6.8) admits a solution. The following methodology shows existence of such a solution and is furthermore crucial for the proof of tightness of $(V^n)_{n\in\mathbb{N}}$. On the one hand it states that under the assumptions on h, δ, f and λ the discrete-time backward stochastic differential equation (6.8) admits a unique solution V^n, which is given by $V^n_{t_k} = v^n_k(S^n_{t_k}, \eta^n_{t_k})$, a fact which is not surprising in view of the Markov structure of (S^n, η^n). On the other hand it also gives a procedure to recursively construct the functions v^n_k, which satisfy certain regularity conditions. The representation $V^n_{t_k} = v^n_k(S^n_{t_k}, \eta^n_{t_k})$ relates to the well-known procedure in the continuous-time case, where the Markov structure of the solution X of some stochastic differential equation $dX_t = f(t, X_{t-})\ dY_t$ yields the existence of a measurable function u, which is the solution of a partial differential equation and which satisfies $X_t = u(t, Y_t)$.

Recall that $V^n_{t_k}$ is recursively *and* implicitly defined by the recursion formula (6.12). This motivates us to define for each $n \in \mathbb{N}$ and $\ell \in I_m$ the function

$$v^{n\ell}_n(x) = h(x, \ell) \tag{6.13}$$

and then for $\ell \in I_m$ and recursively for $k \in \{1, \ldots, n\}$ the following functions:

$$\begin{cases} g^{n\ell}_k(x) &= \sum_{j\in I_m} E_{P'^n}\left[v^{nj}_k(x + \Sigma^\ell(t_k, x)\ \xi^n_k)\right] p^{n\ell j}_k(x) \\ F^{n\ell}_k(x, v) &= g^{n\ell}_k(x) + \gamma^{n\ell}(t_k, x, v)\ \Delta t_k - v \\ v^{n\ell}_{k-1}(x) & \text{such that } F^{n\ell}_k(x, v^{n\ell}_{k-1}(x)) = 0. \end{cases} \tag{6.14}$$

Proposition 6.4 *For n sufficiently large the recursion (6.13)–(6.14) admits a unique solution $(v^{n\ell}_k)_{\ell\in I_m, k\in\{0,\ldots,n\}}$. Moreover all $v^{n\ell}_k$ are C^1-functions on $\mathbb{R}^d$ with $|v^{n\ell}_k(x)|$ and $|\nabla v^{n\ell}_k(x)|$ bounded by some constant c_1, uniformly in n, ℓ, k and x.*

Proof. We claim that the existence and uniqueness of $v^{n\ell}_k$ as well as the boundedness of its gradient follow (recursively) from the implicit function theorem. The boundedness of $v^{n\ell}_k$ will come as a by-product of the recursion formula (6.13)–(6.14), basically from the fact that $F^{n\ell}_k(x, v^{n\ell}_{k-1}(x)) = 0$.

We denote for $y \in \mathbb{R}^d$ the maximum norm of y by $|y|_{\max} = \max_{1\le i\le d} |y^i|$. Note that on $\mathbb{R}^d$ all norms are equivalent; so in order to show uniform boundedness of $\nabla v^{n\ell}_k(x)$ it suffices to show that $|\nabla v^{n\ell}_k(x)|_{\max}$ is uniformly bounded.

For $k = n$ we have $v^{n\ell}_n = h^\ell$, so there is by assumption nothing to show. Then suppose that for some $k \le n$ there exist C^1-functions v^{nj}_k, $j \in I_m$, all bounded by the constant $\bar{c}_k$, with $\left|\nabla v^{nj}_k(x)\right|_{\max} \le \tilde{c}_k$, where we choose the constant $\tilde{c}_k \ge 1$, uniformly in ℓ.

Now we fix $\ell \in I_m$ and define for $j \in I_m$ the functions $g_k^{n\ell j}(x) := E_{P'^n}\left[v_k^{nj}(x + \Sigma^\ell(t_k, x)\ \xi_k^n)\right]$. These are C^1 by Lemma C.7 and bounded by $\bar{c}_k$ by the induction hypothesis, and we have

$$g_k^{n\ell}(x) = \sum_{j \in I_m} g_k^{n\ell j}(x) p_k^{n\ell j}(x).$$

Since $p_k^{n\ell j}$ is bounded and C^1 for all $j \in I_m$ by Lemma 5.14 c), we see that $g_k^{n\ell}$ is bounded and C^1 as well, and by the product rule we have

$$\nabla g_k^{n\ell}(x) = \sum_{j \in I_m} \left(g_k^{n\ell j}(x) \nabla p_k^{n\ell j}(x) + p_k^{n\ell j}(x) \nabla g_k^{n\ell j}(x) \right).$$

Now by construction $|g_k^{n\ell j}| \leq \bar{c}_k$, and $\left|\nabla g_k^{n\ell j}\right|_{\max} \leq \tilde{c}_k \sqrt{1 + \frac{\bar{\Sigma}^2 T}{n}}$ by Lemma C.8. Furthermore $|\nabla p_k^{n\ell j}| \leq \frac{c}{n}$ for some constant c and $\sum_{j \in I_m} p_k^{n\ell j}(x) = 1$ by Lemma 5.14 a). This altogether yields

$$\left|\nabla g_k^{n\ell}\right|_{\max} \leq \bar{c}_k \frac{c}{n} + \tilde{c}_k \sqrt{1 + \frac{\bar{\Sigma}^2 T}{n}}. \tag{6.15}$$

For fixed $x \in \mathbb{R}^d$ we define the function $G_k^{n\ell}(v) := g_k^{n\ell}(x) + \gamma^{n\ell}(t_k, x, v)\ \Delta t_k$. Then since $\gamma^{n\ell}$ is bounded by assumption, we have that $G_k^{n\ell}$ is a contraction (at least for n sufficiently large, hence Δt_k sufficiently small), so that for every $x \in \mathbb{R}^d$ the function $G_k^{n\ell}$ admits a unique fixed point $v \in \mathbb{R}$ by the Banach fixed point theorem. Thus for all $x_0 \in \mathbb{R}^d$ there exists a unique $v_0 \in \mathbb{R}$ such that $F_k^{n\ell}(x_0, v_0) = 0$.

We now apply the implicit function theorem. We know that $g_k^{n\ell}$ and $\gamma^{n\ell}$ are C^1, so $\nabla F_k^{n\ell}$ exists, and since $\frac{\partial}{\partial v}\gamma^{n\ell}$ is bounded we have that $\frac{\partial}{\partial v} F_k^{n\ell}(x, v) = \frac{\partial}{\partial v}\gamma^{n\ell}(t_k, x, v)\ \Delta t_k - 1 \neq 0$ for n sufficiently large. So by the implicit function theorem there exists a unique C^1-function $v_{k-1}^{n\ell} \colon \mathbb{R}^d \to \mathbb{R}$ such that $F_k^{n\ell}(x, v_{k-1}^{n\ell}(x)) = 0$, and whose gradient is given by

$$\begin{aligned} \nabla v_{k-1}^{n\ell}(x) &= -\left(\frac{\partial F_k^{n\ell}}{\partial v}(x, v)\right)^{-1} \nabla_x F_k^{n\ell}(x, v)\Bigg|_{v = v_{k-1}^{n\ell}(x)} \\ &= \frac{\nabla g_k^{n\ell}(x) + \nabla_x \gamma^{n\ell}(t_k, x, v)\ \Delta t_k}{1 - \frac{\partial \gamma^{n\ell}}{\partial v}(t_k, x, v)\ \Delta t_k}\Bigg|_{v = v_{k-1}^{n\ell}(x)}. \end{aligned} \tag{6.16}$$

We first show boundedness of $v_{k-1}^{n\ell}$. Recall that $\left|g_k^{n\ell}(x)\right| \leq \bar{c}_k$, so the construction of $v_{k-1}^{n\ell}$ via $F_k^{n\ell}(x, v_{k-1}^{n\ell}(x)) = 0$ yields

$$|v_{k-1}^{n\ell}(x)| \leq |g_k^{n\ell}(x)| + |\gamma^{n\ell}(t_k, x, v_{k-1}^{n\ell}(x))\ \Delta t_k| \leq \bar{c}_k + c_0 \frac{T}{n} =: \bar{c}_{k-1},$$

so that by backward induction and the fact that $\bar{c}_n = c_0$ (recall that $|v_n^{n\ell}(x)| = |h^\ell(x)| \leq c_0$), we have

$$\bar{c}_{k-1} = c_0 \left(1 + \frac{(n - (k-1))T}{n}\right) \leq c_0(1 + T) =: c_1'$$

for all $k \in \{1, \dots, n\}$, and in particular $|v_{k-1}^{n\ell}(x)| \le c_1'$. Note that for the boundedness of the gradient of $g_k^{n\ell}$ in (6.15), we can therefore assume $\left|\nabla g_k^{n\ell}\right|_{\max} \le c_1' \frac{c}{n} + \tilde{c}_k \sqrt{1 + \frac{\bar{\Sigma}^2 T}{n}}$.

Let us now show the boundedness of the gradient of $v_{k-1}^{n\ell}$. Recall that $|\nabla \gamma^{n\ell}|_{\max} \le c_0$. Then (6.16) yields that for all $\ell \in I_m$

$$
\begin{aligned}
\left|\nabla v_{k-1}^{n\ell}(x)\right|_{\max} &\le \left. \frac{\left|\nabla g_k^{n\ell}(x)\right|_{\max} + \left|\nabla_x \gamma^{n\ell}(t_k, x, v)\right|_{\max} \ \Delta t_k}{\left|1 - \frac{\partial \gamma^{n\ell}}{\partial v}(t_k, x, v) \ \Delta t_k\right|} \right|_{v = v_{k-1}^{n\ell}(x)} \\
&\le \left. \frac{\tilde{c}_k \sqrt{1 + \frac{\bar{\Sigma}^2 T}{n}} + c_1' \frac{c}{n} + c_0 \frac{T}{n}}{1 - \left|\frac{\partial \gamma^{n\ell}}{\partial v}(t_k, x, v)\right| \ \frac{T}{n}} \right|_{v = v_{k-1}^{n\ell}(x)} \\
&\le \frac{\tilde{c}_k \left(1 + \frac{\bar{\Sigma}^2 T}{n}\right) + c_1' \frac{c}{n} + c_0 \frac{T}{n}}{1 - c_0 \frac{T}{n}} \\
&\le \tilde{c}_k \left(1 + \frac{\frac{(\bar{\Sigma}^2 + 2c_0)T + c_1' c}{n}}{1 - c_0 \frac{T}{n}}\right) \\
&\le \tilde{c}_k \left(1 + \frac{2(\bar{\Sigma}^2 + 2c_0)T + c_1' c}{n}\right),
\end{aligned}
$$

where in the fourth inequality we have used $\tilde{c}_k \ge 1$, and in the last inequality $1 - c_0 \frac{T}{n} \ge \frac{1}{2}$ for n sufficiently large. Now by assumption we have $\tilde{c}_n = c_0$ (recall $|\nabla v_n^{n\ell}(x)| = |\nabla h^\ell(x)| \le c_0$ by assumption), so by backward induction we have

$$
\left|\nabla v_{k-1}^{n\ell}(x)\right|_{\max} \le c_0 \left(1 + \frac{2(\bar{\Sigma}^2 + 2c_0)T + c_1' c}{n}\right)^{n-(k-1)} \le c_0 \ e^{2(\bar{\Sigma}^2 + 2c_0)T + c_1' c}
$$

for n sufficiently large. Finally since all norms on $\mathbb{R}^d$ are equivalent, we get

$$
\left|\nabla v_{k-1}^{n\ell}(x)\right| \le c_{\max} \left|\nabla v_{k-1}^{n\ell}(x)\right|_{\max} \le c_{\max} \ c_0 \ e^{2(\bar{\Sigma}^2 + 2c_0)T + c_1' c} =: c_1''
$$

for some constant $c_{\max}$. At last we set $c_1 = c_1' \vee c_1''$. □

We are now in a position to show the existence of a solution V^n to (6.8). Recall that uniqueness of a solution was shown in Proposition 6.3.

Proposition 6.5 *Let $n \in \mathbb{N}$ be sufficiently large and let $(v_k^{n\ell})_{\ell,k}$ be the unique solution of (6.13)–(6.14). Then the piecewise constant càdlàg process V^n, defined by*

$$
V_{t_k}^n := v_k^n(S_{t_k}^n, \eta_{t_k}^n) := \sum_{\ell \in I_m} v_k^{n\ell}(S_{t_k}^n) \ \mathbb{1}_{\{\eta_{t_k}^n = \ell\}}, \tag{6.17}
$$

solves the backward stochastic differential equation (6.8).

Proof. By Lemma 6.2 it suffices to show that V^n given by (6.17) solves the recursion (6.12). The definition of $v_n^{n\ell}$ yields $V_T^n = h(S_T^n, \eta_T^n)$, so by backward induction it remains to show that for $k \in \{0, \dots, n-1\}$ we have

$$v_k^n(S_{t_k}^n, \eta_{t_k}^n) = E_{P^n}\left[v_{k+1}^n(S_{t_{k+1}}^n, \eta_{t_{k+1}}^n)\middle|\, \mathcal{F}_{t_k}^n\right] + \gamma^n\left(t_{k+1}, S_{t_k}^n, \eta_{t_k}^n, v_k^n(S_{t_k}^n, \eta_{t_k}^n)\right)\Delta t_k. \tag{6.18}$$

For $\ell \in I_m$ we have $v_k^n(x,\ell) = v_k^{n\ell}(x)$, where $v_k^{n\ell}$ satisfies by definition

$$v_k^{n\ell}(x) = g_{k+1}^{n\ell}(x) + \gamma^{n\ell}(t_{k+1}, x, v_k^{n\ell}(x))\ \Delta t_{k+1}$$

with $g_{k+1}^{n\ell}$ as defined in (6.14). Note that by Lemma 5.15 with $\Gamma \equiv 0$ we have that

$$\begin{aligned}
&E_{P^n}\left[v_{k+1}^n(S_{t_{k+1}}^n, \eta_{t_{k+1}}^n)\middle|\, \mathcal{F}_{t_k}^n\right] \\
&= \sum_{\ell, j \in I_m} E_{P'^n}\left[v_{k+1}^n(x + \Sigma(t_{k+1}, x, \ell)\xi_{k+1}^n, j)\right]\Big|_{x=S_{t_k}^n}\ p_{k+1}^{n\ell j}(S_{t_k}^n)\ \mathbb{1}_{\{\eta_{t_k}^n = \ell\}} \\
&= \sum_{\ell \in I_m} g_{k+1}^{n\ell}(S_{t_k}^n)\ \mathbb{1}_{\{\eta_{t_k}^n = \ell\}},
\end{aligned}$$

so that on $\{\eta_{t_k}^n = \ell\}$ we have

$$\begin{aligned}
v_k^n(S_{t_k}^n, \eta_{t_k}^n) &= v_k^{n\ell}(S_{t_k}^n) \\
&= g_{k+1}^{n\ell}(S_{t_k}^n) + \gamma^{n\ell}(t_{k+1}, S_{t_k}^n, v_k^{n\ell}(S_{t_k}^n))\ \Delta t_{k+1} \\
&= E_{P^n}\left[v_{k+1}^n(S_{t_{k+1}}^n, \eta_{t_{k+1}}^n)\middle|\, \mathcal{F}_{t_k}^n\right] + \gamma^{n\ell}(t_{k+1}, S_{t_k}^n, v_k^{n\ell}(S_{t_k}^n))\ \Delta t_{k+1},
\end{aligned}$$

which shows (6.18). □

6.3 Tightness of the Sequence of Approximating Price Processes

We now show tightness of $(\mathcal{L}(V^n|P'^n))_{n\in\mathbb{N}}$. At first sight it might seem more natural to show tightness of $(V^n)_{n\in\mathbb{N}}$ under $(P^n)_{n\in\mathbb{N}}$, since V^n is defined via conditional expectations under P^n, and indeed with the results of Section 5.4 the following arguments extend easily to show tightness of $(\mathcal{L}(V^n|P^n))_{n\in\mathbb{N}}$. But in Section 6.4 it will be important that the filtrations under consideration are generated by processes with independent increments, which is the case under P'^n but not under P^n.

The idea to prove tightness is to apply Theorem 1.45, so that we need to show that for the canonical decomposition $V^n = V_0^n + M^n + A^n$ the sequence $(V_0^n)_{n\in\mathbb{N}}$ is tight and that for each $n \in \mathbb{N}$ the increasing process

$$\hat{G}^n = \langle M^n \rangle + \mathrm{Var}(A^n)$$

is strongly dominated by a predictable process G^n such that the sequence $(G^n)_{n\in\mathbb{N}}$ converges weakly to a continuous process. (Recall that $\mathrm{Var}(A^n) = \int |dA^n|$ denotes the total variation process of A^n.)

Theorem 6.6 *Let V^n be given by (6.8) or (6.12). Suppose h^ℓ, δ^ℓ, f^{ij} and λ^{ij} are bounded and C^1 for all $\ell \in I_m$, $i, j \in \{1, \dots, m\}$, with gradients bounded uniformly in x and v. If furthermore all partial derivatives $\frac{\partial(\Sigma^\ell)^{ij}}{\partial x^k}$ for $\ell \in I_m$, $i, k \in \{1, \dots, d\}$, $j \in \{1, \dots, r\}$ are bounded, uniformly in t, then the sequence $(\mathcal{L}(V^n | P'^n))_{n \in \mathbb{N}}$ is tight.*

The proof of Theorem 6.6 is divided into several steps, and we start with the following remarks. It is well-known that the Doob decomposition $V^n = V_0^n + M^n + A^n$ under P'^n of the discrete-time process V^n is given by $M_0^n = A_0^n = 0$,

$$\text{(6.19)} \qquad \begin{cases} \Delta M_{t_k}^n &= V_{t_k}^n - E_{P'^n}\left[V_{t_k}^n \middle| \mathcal{F}_{t_{k-1}}^n \right], \\ \Delta A_{t_k}^n &= E_{P'^n}\left[V_{t_k}^n - V_{t_{k-1}}^n \middle| \mathcal{F}_{t_{k-1}}^n \right], \end{cases}$$

$k \in \{1, \dots, n\}$, and piecewise constant càdlàg interpolation between t_k and t_{k+1}. We have shown in Proposition 6.5 that $V_{t_k}^n = v_k^n(S_{t_k}^n, \eta_{t_k}^n)$ for every k. This allows us to handle the processes $\langle M^n \rangle$ and $\mathrm{Var}(A^n)$. Note that once the functions v_k^n from Proposition 6.5 are fixed, the representation $V_{t_k}^n = v_k^n(S_{t_k}^n, \eta_{t_k}^n)$ does not depend on the probability measure.

Recall that we write $\langle\!\langle X \rangle\!\rangle := \sum_{i=1}^d \langle X^i \rangle$ for a d-dimensional local martingale X for which each component is locally square-integrable.

Proposition 6.7 *There exists a constant $c_2 < \infty$ such that the increasing process G^n, given by*

$$G^n = c_2 \left(\langle\!\langle S^n \rangle\!\rangle + L^n \right),$$

strongly dominates $\langle M^n \rangle + \mathrm{Var}(A^n)$.

Proof. 1) We show $|\Delta A_{t_k}^n| \le c\frac{T}{n} = c\Delta L_{t_k}^n$, where c is the constant given in Lemma 5.14, so that $\mathrm{Var}(A^n) \prec cL^n =: G^{n,A}$. Recall that $t_k = k\frac{T}{n}$ and $n(t) = \left[\frac{nt}{T}\right]$. By the recursion formula (6.12) we have

$$V_{t_{k-1}}^n = E_{P^n}\left[V_{t_k}^n \middle| \mathcal{F}_{t_{k-1}}^n \right] + \gamma^n(t_k, S_{t_{k-1}}^n, \eta_{t_{k-1}}^n, V_{t_{k-1}}^n)\ \Delta t_k,$$

and Lemma 5.15 yields

$$\begin{aligned} E_{P^n}\left[V_{t_k}^n \middle| \mathcal{F}_{t_{k-1}}^n \right] &= E_{P^n}\left[v_k^n(S_{t_k}^n, \eta_{t_k}^n) \middle| \mathcal{F}_{t_{k-1}}^n \right] \\ &= \sum_{\ell, j \in I_m} E_{P'^n}\left[v_k^n(S_{t_k}^n, j) \middle| \mathcal{F}_{t_{k-1}}^n \right]\ p_k^{n\ell j}(S_{t_{k-1}}^n)\ \mathbb{1}_{\{\eta_{t_{k-1}}^n = \ell\}}. \end{aligned}$$

On the other hand Lemma 5.15 together with Remark 5.16 shows that

$$E_{P'^n}\left[V_{t_k}^n \middle| \mathcal{F}_{t_{k-1}}^n \right] = \sum_{\ell, j \in I_m} E_{P'^n}\left[v_k^n(S_{t_k}^n, j) \middle| \mathcal{F}_{t_{k-1}}^n \right]\ p'^{n\ell j}\ \mathbb{1}_{\{\eta_{t_{k-1}}^n = \ell\}},$$

so that

$$\begin{aligned}\Delta A^n_{t_k} &= E_{P'^n}\left[V^n_{t_k} - V^n_{t_{k-1}}\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\\ &= \sum_{\ell,j\in I_m} E_{P'^n}\left[v^n_k(S^n_{t_k},j)\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\left(p'^{n\ell j} - p^{n\ell j}_k(S^n_{t_{k-1}})\right)\mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}}\\ &\qquad -\gamma^n(t_k, S^n_{t_{k-1}}, \eta^n_{t_{k-1}}, V^n_{t_{k-1}})\;\Delta t_k.\end{aligned}$$

Now v^n_k and γ^n are bounded, uniformly in n, whereas $|p'^{n\ell j} - p^{n\ell j}_k(x)| \le 2\frac{c}{n}$. To see this, recall that $0 \le p'^{n\ell j}, p^{n\ell j}_k(x) \le \frac{c}{n}$ for $\ell \ne j$ by Lemma 5.14 and Remark 5.16, whereas for $\ell = j$ we have $1 - \frac{c}{n} \le p'^{n\ell\ell}, p^{n\ell\ell}_k(x) \le 1$, so that $-\frac{c}{n} \le p'^{n\ell\ell} - p^{n\ell\ell}_k(x) \le \frac{c}{n}$. Altogether this yields

$$|\Delta A^n_{t_k}| \le c\,\frac{T}{n} = c\;\Delta L^n_{t_k}.$$

2) Let us turn to the quadratic variation of M^n, the martingale part in the Doob decomposition of V^n. Recall that for a locally square-integrable martingale X in the filtration $\mathbb{F}^n$ we have $\Delta\langle X\rangle_{t_k} = E\big[(X_{t_k} - X_{t_{k-1}})^2\big|\mathcal{F}^n_{t_{k-1}}\big]$ and $\langle X\rangle_t = \langle X\rangle_{t_k}$ for $t_k \le t < t_{k+1}$. We show that for all $k \in \{1,\dots,n\}$ we have $\langle M^n\rangle_{t_k} - \langle M^n\rangle_{t_{k-1}} \le G^{n,M}_{t_k} - G^{n,M}_{t_{k-1}}$, for a suitable predictable process $G^{n,M}$. More precisely we show

$$\begin{aligned}E_{P'^n}\left[\left(V^n_{t_k} - E_{P'^n}\left[V^n_{t_k}\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right] &\le \bar{c}\left(E_{P'^n}\left[\left|S^n_{t_k} - S^n_{t_{k-1}}\right|^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right] + \frac{1}{n}\right)\\ &= \bar{c}\left(\Delta\langle\!\langle S^n\rangle\!\rangle_{t_k} + \frac{1}{n}\right)\end{aligned}$$

for some $\bar{c} < \infty$. The assertion of Proposition 6.7 then follows by summing over k and setting $c_2 = T\bar{c} + c$, so that $\langle M^n\rangle + \mathrm{Var}(A^n) \prec G^{n,M} + G^{n,A} \prec G^n$.

To start with, it is clear that

$$E_{P'^n}\left[\left(V^n_{t_k} - E_{P'^n}\left[V^n_{t_k}\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right] = E_{P'^n}\left[\left(V^n_{t_k}\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right] - \left(E_{P'^n}\left[V^n_{t_k}\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\right)^2,$$

and by Proposition 6.5 we have $V^n_{t_k} = v^n_k(S^n_{t_k}, \eta^n_{t_k}) = \sum_{j\in I_m} v^{nj}_k(S^n_{t_k})\;\mathbb{1}_{\{\eta^n_{t_k}=j\}}$ for bounded C^1-functions v^{nj}_k so that

$$\begin{aligned}E_{P'^n}\left[\left(V^n_{t_k}\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right] &= E_{P'^n}\left[\left(v^n_k(S^n_{t_k},\eta^n_{t_k})\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\\ &= \sum_{\ell,j\in I_m} E_{P'^n}\left[\left(v^n_k(S^n_{t_k},j)\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\;p'^{n\ell j}\;\mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}}\\ &= \sum_{\ell,j\in I_m} E_{P'^n}\left[\left(v^{nj}_k(S^n_{t_k})\right)^2\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\;p'^{n\ell j}\;\mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}},\end{aligned}$$

by applying Lemma 5.15 and Remark 5.16, and with the same reasoning

$$E_{P'^n}\left[V^n_{t_k}\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right] = \sum_{\ell,j\in I_m} E_{P'^n}\left[v^{nj}_k(S^n_{t_k})\,\middle|\,\mathcal{F}^n_{t_{k-1}}\right]\;p'^{n\ell j}\;\mathbb{1}_{\{\eta^n_{t_{k-1}}=\ell\}}.$$

With this we have on $\{\eta^n_{t_{k-1}} = \ell\}$ for $\ell \in I_m$

$$
\begin{aligned}
& E_{P'^n}\left[\left(V^n_{t_k}\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] - \left(E_{P'^n}\left[V^n_{t_k} \middle| \mathcal{F}^n_{t_{k-1}}\right]\right)^2 \\
= & \sum_{j \in I_m} E_{P'^n}\left[\left(v^{nj}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell j} - \left(\sum_{j \in I_m} E_{P'^n}\left[v^{nj}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell j}\right)^2 \\
= & \sum_{\substack{j \in I_m \\ j \neq \ell}} \left(E_{P'^n}\left[\left(v^{nj}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell j} - \left(E_{P'^n}\left[v^{nj}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell j}\right)^2\right) \\
& \quad - \sum_{\substack{j_1, j_2 \in I_m \\ j_1 \neq j_2}} E_{P'^n}\left[v^{nj_1}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] E_{P'^n}\left[v^{nj_2}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell j_1} \, p'^{n\ell j_2} \\
& \quad + E_{P'^n}\left[\left(v^{n\ell}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell\ell} - \left(E_{P'^n}\left[v^{n\ell}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell\ell}\right)^2 .
\end{aligned}
$$

Now recall Lemma 5.14 and Proposition 6.4. For $\ell \neq j$ we have $p'^{n\ell j} \leq \frac{c}{n}$, and $|v^{n\ell}_k(x)| \leq c_1$, uniformly in n, ℓ, k and x, so that we get on $\{\eta^n_{t_{k-1}} = \ell\}$

$$
\begin{aligned}
& \left| E_{P'^n}\left[\left(V^n_{t_k}\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] - \left(E_{P'^n}\left[V^n_{t_k} \middle| \mathcal{F}^n_{t_{k-1}}\right]\right)^2 \right| \\
\leq & \; \tilde{c} \, \frac{1}{n} + \left| E_{P'^n}\left[\left(v^{n\ell}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell\ell} - \left(E_{P'^n}\left[v^{n\ell}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell\ell}\right)^2 \right| ,
\end{aligned}
$$

for $\tilde{c} = m(m+3)c_1^2 < \infty$. Concerning an upper bound for the remaining term above, we first note that

$$
\begin{aligned}
& \left| E_{P'^n}\left[\left(v^{n\ell}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell\ell} - \left(E_{P'^n}\left[v^{n\ell}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right] \, p'^{n\ell\ell}\right)^2 \right| \\
\leq & \left| E_{P'^n}\left[\left(v^{n\ell}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] - \left(E_{P'^n}\left[v^{n\ell}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right]\right)^2 \right| \left| p'^{n\ell\ell} \right|^2 \\
& \quad + E_{P'^n}\left[\left(v^{n\ell}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] p'^{n\ell\ell} \left(1 - p'^{n\ell\ell}\right) \\
\leq & \left| E_{P'^n}\left[\left(v^{n\ell}_k(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] - \left(E_{P'^n}\left[v^{n\ell}_k(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right]\right)^2 \right| + \frac{c_1^2 c}{n},
\end{aligned}
$$

since $|v^{n\ell}_k| \leq c_1$ by Proposition 6.4, and since $p'^{n\ell\ell} \leq 1$ and $1 - p'^{n\ell\ell} \leq \frac{c}{n}$ by Lemma 5.14.

Now recall that the $v^{n\ell}_k$ are C^1 with $|\nabla v^{n\ell}_k|$ uniformly bounded by c_1 from Proposition 6.4, so we have by a Taylor expansion

$$
\begin{aligned}
v^{n\ell}_k(S^n_{t_k}) &= v^{n\ell}_k\left(S^n_{t_{k-1}} + (S^n_{t_k} - S^n_{t_{k-1}})\right) \\
&= v^{n\ell}_k(S^n_{t_{k-1}}) + \nabla v^{n\ell}_k\left(S^n_{t_{k-1}} + \vartheta \, (S^n_{t_k} - S^n_{t_{k-1}})\right)(S^n_{t_k} - S^n_{t_{k-1}}),
\end{aligned}
$$

for some (random) $\vartheta \in [0,1]$. Now since $v_k^{n\ell}(S^n_{t_{k-1}})$ is $\mathcal{F}^n_{t_{k-1}}$-measurable, we get

$$\begin{aligned}
& E_{P'^n}\left[\left(v_k^{n\ell}(S^n_{t_k})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] - \left(E_{P'^n}\left[v_k^{n\ell}(S^n_{t_k}) \middle| \mathcal{F}^n_{t_{k-1}}\right]\right)^2 \\
= \;& E_{P'^n}\left[\left(\nabla v_k^{n\ell}\left(S^n_{t_{k-1}} + \vartheta\,(S^n_{t_k} - S^n_{t_{k-1}})\right)(S^n_{t_k} - S^n_{t_{k-1}})\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \\
& - \left(E_{P'^n}\left[\nabla v_k^{n\ell}\left(S^n_{t_{k-1}} + \vartheta\,(S^n_{t_k} - S^n_{t_{k-1}})\right)(S^n_{t_k} - S^n_{t_{k-1}}) \middle| \mathcal{F}^n_{t_{k-1}}\right]\right)^2 \\
\leq \;& c_1\, E_{P'^n}\left[\left|S^n_{t_k} - S^n_{t_{k-1}}\right|^2 \middle| \mathcal{F}^n_{t_{k-1}}\right]
\end{aligned}$$

by Lemma C.9. Putting these results together we obtain for $\bar{c}$ chosen sufficiently large

$$E_{P'^n}\left[\left(V^n_{t_k} - E_{P'^n}\left[V^n_{t_k}\middle|\mathcal{F}^n_{t_{k-1}}\right]\right)^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] \leq \bar{c}\,\left(E_{P'^n}\left[\left|S^n_{t_k} - S^n_{t_{k-1}}\right|^2 \middle| \mathcal{F}^n_{t_{k-1}}\right] + \frac{1}{n}\right),$$

so we obtain the desired inequality. □

In order to prove Theorem 6.6 it remains to show that $G^n = c_2\,(\langle\!\langle S^n \rangle\!\rangle + L^n)$ converges in distribution to a continuous process and that $(\mathcal{L}(V_0^n)|P'^n)_{n\in\mathbb{N}}$ is tight. Notice that L^n is deterministic and converges uniformly on compacts to the deterministic continuous process L with $L_t = t$, so it is sufficient to show that $\langle\!\langle S^n \rangle\!\rangle$ converges to a continuous process.

Lemma 6.8 *Let S^n and S be given by (6.6) and (6.1). Then $\mathcal{L}(\langle\!\langle S^n \rangle\!\rangle|P'^n) \xrightarrow{w} \mathcal{L}(\langle\!\langle S \rangle\!\rangle|P')$.*

Proof. Since S^n satisfies the stochastic differential equation (6.6), and since the driving processes W^{nj}, $j \in \{1,\dots,r\}$ are independent under P'^n and thus orthogonal, we see that

$$\begin{aligned}
\langle\!\langle S^n \rangle\!\rangle_t &= \sum_{i=1}^d \left\langle \sum_{j=1}^r \int \Sigma^{ij}(L^n_{s-}, S^n_{s-}, \eta^n_{s-})\; dW^{nj}_s \right\rangle_t \\
&= \sum_{i=1}^d \sum_{j=1}^r \left\langle \int \Sigma^{ij}(L^n_{s-}, S^n_{s-}, \eta^n_{s-})\; dW^{nj}_s \right\rangle_t \\
&= \sum_{i=1}^d \sum_{j=1}^r \int_0^t \left(\Sigma^{ij}(L^n_{s-}, S^n_{s-}, \eta^n_{s-})\right)^2\; d\left\langle W^{nj}\right\rangle_s \\
&= \sum_{i=1}^d \left(\int_0^t H(L^n_{s-}, S^n_{s-}, \eta^n_{s-})\; d\mathbb{L}^n_s\right)^i
\end{aligned}$$

where H is a matrix-valued function with $H^{ij}(s,x,y) = (\Sigma^{ij}(s,x,y))^2$, and $\mathbb{L}^n$ is an r-dimensional process with $\mathbb{L}^{nj}_s = \left\langle W^{nj}\right\rangle_s = L^n_s$ for all $j \in \{1,\dots,r\}$. With the same argument and notation we have

$$\langle\!\langle S \rangle\!\rangle_t = \sum_{i=1}^d \left(\int_0^t H(s, S_s, \eta_{s-})\; d\mathbb{L}_s\right)^i,$$

for $\mathbb{L}_s^j = s$. In order to show convergence we note that by Theorem 5.6 we have that $\mathcal{L}(L^n, S^n, \eta^n | P'^n) \xrightarrow{w} \mathcal{L}(L, S, \eta | P')$. If we consider on $\mathbb{R}^{d+1} \times I_m$ the product topology of the usual topology on $\mathbb{R}^{d+1}$ and the discrete topology on I_m, then $H(s, x, y)$ is continuous in (s, x) by assumption and trivially in y. Since we consider the discrete topology in the y-argument, we also have that H is continuous in (s, x, y). Therefore by the continuous mapping theorem we get

$$U^n := H(L^n, S^n, \eta^n) \xrightarrow{\mathcal{L}} H(L, S, \eta) =: U$$

under P'^n, P', and since $\mathbb{L}^n$ and $\mathbb{L}$ are deterministic processes we also get joint convergence $\mathcal{L}(U^n, \mathbb{L}^n | P'^n) \xrightarrow{w} (U, \mathbb{L} | P')$. Now by Example 1.51 $(\mathbb{L}^n)_{n\in\mathbb{N}}$ is good, so we get by Definition 1.48

$$\int U_-^n \, d\mathbb{L}^n \xrightarrow{\mathcal{L}} \int U_- d\mathbb{L}$$

under P'^n, P', which implies joint convergence of the coordinates of $\int U_-^n \, d\mathbb{L}^n$, and therefore

$$\langle\!\langle S^n \rangle\!\rangle = \sum_{i=1}^d \left(\int H(L_-^n, S_-^n, \eta_-^n) \, d\mathbb{L}^n \right)^i \xrightarrow{\mathcal{L}} \sum_{i=1}^d \left(\int_0^t H(L, S, \eta_-) \, d\mathbb{L} \right)^i = \langle\!\langle S \rangle\!\rangle$$

under P'^n, P', so the proof is finished. □

Lemma 6.9 *The sequence $(\mathcal{L}(V_0^n) | P'^n)_{n\in\mathbb{N}}$ is tight.*

Proof. Note that by definition V_0^n is $\mathcal{F}_0^n$-measurable and thus deterministic for all $n \in \mathbb{N}$. Thus it suffices to show that $(V_0^n)_{n\in\mathbb{N}}$ is a bounded sequence of real numbers. However by definition we have

$$V_0^n \quad = \quad E_{P^n}\left[h(S_T^n, \eta_T^n) + \int_0^T \delta(s, S_{s-}^n, \eta_{s-}^n, V_{s-}^n) \, dL_s^n + \int_0^T \sum_{\ell,j=1}^m f^{\ell j}(s, S_{s-}^n, V_{s-}^n) \, dN_s^{n\ell j} \right],$$

and since h, δ and $f^{\ell j}$ are bounded there exists a constant c such that for all $n \in \mathbb{N}$

$$|V_0^n| \le c \left(1 + \sum_{\ell,j=1}^m E_{P^n}\left[N_T^{n\ell j} \right] \right).$$

Now $N_T^{n\ell j} = \sum_{k=1}^n \zeta_k^{n\ell j}$, so we get by Lemma 5.8

$$E_{P^n}\left[N_T^{n\ell j} \right] = \sum_{k=1}^n E_{P^n}\left[\zeta_k^{n\ell j} \right] = \sum_{k=1}^n E_{P^n}\left[E_{P^n}\left[\zeta_k^{n\ell j} \middle| \mathcal{F}_{t_{k-1}}^n \right] \right] = \sum_{k=1}^n E_{P^n}\left[\bar{\lambda}^{n\ell j}(t_k, S_{t_{k-1}}^n) \right] \Delta t_k,$$

for $\bar{\lambda}^{n\ell j} = \lambda^{\ell j} + \frac{T}{n}\left(1 - \lambda^{\ell j}\right)$. Now the uniform boundedness by c_0 of $\lambda^{\ell j}$ yields that $\bar{\lambda}^{n\ell j}$ is bounded by some constant $\tilde{c}_0$, and with $\Delta t_k = \frac{T}{n}$ we get $E_{P^n}\left[N_T^{n\ell j}\right] \le \tilde{c}_0 T$, and therefore $|V_0^n| \le c\left(1 + m^2 \tilde{c}_0 T\right)$, which finishes the proof. □

Proof of Theorem 6.6. By Proposition 6.7 we have $\langle M^n\rangle + \mathrm{Var}(A^n) \prec G^n$ for the predictable process $G^n = c_2\left(\langle\!\langle S^n\rangle\!\rangle + L^n\right)$. $(L^n)_{n\in\mathbb{N}}$ is a sequence of deterministic processes converging to the deterministic process L with $L_t = t$, so this together with Lemma 6.8 yields that G^n converges in distribution to the continuous process G with $G_t = c_2\left(\langle\!\langle S\rangle\!\rangle_t + t\right)$. Finally $(\mathcal{L}(V_0^n|P'^n))_{n\in\mathbb{N}}$ is tight by Lemma 6.9, so $(\mathcal{L}(V^n|P'^n))_{n\in\mathbb{N}}$ is tight by Theorem 1.45. □

6.4 Convergence of Price Processes

Our aim is of course to show convergence of $(V^n)_{n\in\mathbb{N}}$ rather than tightness only. As usual the idea is to identify every cluster point of $(V^n)_{n\in\mathbb{N}}$ with V. However, V^n and V are both defined via conditional expectations with respect to some filtration, and convergence in distribution has at first nothing to do with filtrations. Therefore we use the Skorokhod embedding theorem to find a new probability space in order to get almost sure convergence of the processes involved, which allows us to use results on the concept of *convergence of filtrations*.

Recall that the Skorokhod topology as explained in Appendix B is metrizable and separable. Thus it makes sense to define for càdlàg processes X^n and X *P-a.s. convergence for the Skorokhod topology* by $P[\{\omega\colon X^n(\omega) \xrightarrow{S} X(\omega)\}] = 1$, where $\alpha^n \xrightarrow{S} \alpha$ denotes convergence of càdlàg functions with respect to the Skorokhod topology, cf. (B.1). In the same vein we define *convergence in probability for the Skorokhod topology*, denoted by $X^n \xrightarrow{P} X$ for the Skorokhod topology, by $\lim_{n\to\infty} P[\{\omega\colon \delta(X^n(\omega), X(\omega)) > \varepsilon\}] = 0$ for all $\varepsilon > 0$, where δ denotes the Skorokhod metric (cf. Jacod and Shiryaev (1987), VI.1.26). Note that $X^n \to X$ P-a.s. for the Skorokhod topology implies $X^n \xrightarrow{P} X$ for the Skorokhod topology, whereas $X^n \xrightarrow{P} X$ for the Skorokhod topology implies that there exists a subsequence X^{n_k} such that $X^{n_k} \xrightarrow{k\to\infty} X$ P-a.s. for the Skorokhod topology. In summary, convergence (P-a.s. or in probability) for the Skorokhod topology is simply the usual concept for random elements $(X^n)_{n\in\mathbb{N}}$ and X with values in the Skorokhod space of càdlàg functions, viewed as metric space with the Skorokhod metric.

For the remaining part of this chapter recall that all filtrations, whether abstract or generated by a stochastic process, are taken to be right-continuous and that all semimartingales are taken to have càdlàg paths. For the following results we denote for an integrable random variable w and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ by $E[w|\mathbb{F}]$ the càdlàg version of the martingale M with $M_t = E[w|\mathcal{F}_t]$, $t \in [0,T]$.

Definition 6.10 Let $\mathbb{F}^n = (\mathcal{F}^n_t)_{t\in[0,T]}$, $n \in \mathbb{N}$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be filtrations on some fixed probability space $(\Omega, \mathcal{F}, P)$. We say that the sequence $(\mathbb{F}^n)_{n\in\mathbb{N}}$ *converges weakly* to $\mathbb{F}$, and we write $\mathbb{F}^n \xrightarrow{w} \mathbb{F}$, if for all $B \in \mathcal{F}_T$ the sequence of càdlàg martingales $E[\mathbb{1}_B|\mathbb{F}^n]$ converges in probability for the Skorokhod topology on $\mathbb{D}(\mathbb{R})$ to the martingale $E[\mathbb{1}_B|\mathbb{F}]$. ◇

Lemma 6.11 *Let X^n be a sequence of càdlàg processes with independent increments. If $X^n \to X$ in probability for the Skorokhod topology on $\mathbb{D}(\mathbb{R}^d)$, then $\mathbb{F}^{X^n} \xrightarrow{w} \mathbb{F}^X$.*

Proof. cf. Coquet, Mémin and Słominski (2001), Proposition 2. □

Lemma 6.12 *Let w^n and w be $\mathbb{R}^d$-valued random variables. If $w^n \to w$ in $\mathcal{L}^1(P)$, then $\mathbb{F}^n \xrightarrow{w} \mathbb{F}$ implies $E[w^n|\mathbb{F}^n] \xrightarrow{P} E[w|\mathbb{F}]$ for the Skorokhod topology on $\mathbb{D}(\mathbb{R}^d)$.*

Proof. See the remark after Lemma 1 in Coquet, Mémin and Słominski (2001). □

In Becherer and Schweizer (2003) it is shown, using the fact that $\int \lambda^{\ell j}(t, S_t)\ dt$ is the P-compensator of $N^{\ell j}$, that if V satisfies (6.4), then V also satisfies

$$V_t = E_P\left[h(S_T,\eta_T) + \int_t^T \gamma(s, S_s, \eta_{s-}, V_{s-})\ ds \middle| \mathcal{F}_t\right], \quad t \in [0,T], \tag{6.20}$$

where

$$\gamma(t,x,\ell,v) := \delta^\ell(t,x,v) + \sum_{i,j=1}^m f^{ij}(t,x,v)\ \lambda^{ij}(t,x),$$

and that the solution of (6.20) is unique in the class of bounded $\mathbb{F}$-semimartingales, cf. Lemma 4.2 of Becherer and Schweizer (2003) and the proof thereof. Now if V satisfies (6.4), it is clearly $\mathbb{F}$-adapted, and thus V also satisfies

$$V_t = E_P\left[h(S_T,\eta_T) + \int_0^T \gamma(s, S_s, \eta_{s-}, V_{s-})\ ds \middle| \mathcal{F}_t\right] - \int_0^t \gamma(s, S_s, \eta_{s-}, V_{s-})\ ds \tag{6.21}$$

for all $t \in [0,T]$. Conversely it is clear that every $\mathbb{F}$-*adapted* process V which satisfies (6.21) also satisfies (6.20) and then also the backward stochastic differential equation (6.4). By the tightness of $(V^n)_{n\in\mathbb{N}}$ and convergence of filtrations we will in the sequel obtain a càdlàg process $\hat{V}$ which solves (6.21) but is a priori merely adapted to a usually bigger filtration. The next results show that such a process is actually $\mathbb{F}$-adapted and thus also satisfies (6.20). The key idea is to show that for a slight variation (6.22) of (6.21), we have uniqueness of a solution even in the class of bounded product-measurable càdlàg processes. To that end we define

$$\mathcal{X}_b = \{X\colon \Omega\times[0,T]\to\mathbb{R},\ X \text{ is a bounded } \mathcal{F}\otimes\mathcal{B}([0,T])\text{-measurable càdlàg process}\}.$$

Lemma 6.13 *Let $M \in \mathcal{X}_b$ be fixed. Then the stochastic differential equation*

$$V_t = M_t - \int_0^t \gamma(s, S_s, \eta_{s-}, V_{s-})\ ds. \tag{6.22}$$

admits at most one solution in $\mathcal{X}_b$.

Proof. The proof is closely related to the proof of Lemma 4.2 of Becherer and Schweizer (2003). Let $V^1, V^2 \in \mathcal{X}_b$ be two solutions of (6.22). Then since γ is globally Lipschitz in the last argument, uniformly in the other arguments, there exists a constant L_γ such that $|\gamma(s,x,y,v^1) - \gamma(s,x,y,v^2)| \leq L_\gamma |v^1 - v^2|$ for all s,x,y,v^1,v^2. So for all $\beta > 0$ and $t \in [0,T]$ we have

$$\begin{aligned} e^{-\beta t}\left|V_t^1 - V_t^2\right| &\leq e^{-\beta t}\int_0^t \left|\gamma(s,S_s,\eta_{s-},V_{s-}^1) - \gamma(s,S_s,\eta_{s-},V_{s-}^2)\right| \, ds \\ &\leq L_\gamma e^{-\beta t}\int_0^t e^{\beta s}\left\| \sup_{0\leq s\leq T} e^{-\beta s}\left|V_s^1 - V_s^2\right| \right\|_{L^\infty} ds \\ &\leq \frac{L_\gamma}{\beta}\left\| \sup_{0\leq s\leq T} e^{-\beta s}\left|V_s^1 - V_s^2\right| \right\|_{L^\infty}. \end{aligned}$$

So taking the supremum over $t \in [0,T]$ of the above inequality we get a contradiction for $\beta > L_\gamma$ if V^1 and V^2 are not indistinguishable. □

Lemma 6.14 *Suppose now that M in Lemma 6.13 is a bounded $(P,\mathbb{F})$-semimartingale. If (6.22) has a solution $V \in \mathcal{X}_b$, then V is a $(P,\mathbb{F})$-semimartingale.*

Proof. Since γ is globally Lipschitz in the last argument, and since S and η are càdlàg $\mathbb{F}$-adapted processes, $f\colon \mathbb{R}_+ \times \Omega \times \mathbb{R} \to \mathbb{R}$, $(t,\omega,v) \mapsto f(t,\omega,v) := \gamma(t,S_t(\omega),\eta_{t-}(\omega),v)$ is random Lipschitz and thus induces a functional Lipschitz operator in the sense of Protter (1990), definitions on pp. 194-195. So (6.22) admits a unique semimartingale solution $\tilde{V}$ by Protter (1990), Theorem V.7, and since γ and M are bounded by assumption, $\tilde{V}$ is bounded. Now every bounded semimartingale is in $\mathcal{X}_b$, and since V is the only solution of (6.22) in $\mathcal{X}_b$ by Lemma 6.13, we have that $\tilde{V}$ and V are indistinguishable, and thus V is a $(P,\mathbb{F})$-semimartingale. □

Corollary 6.15 *Let $\hat{V} \in \mathcal{X}_b$ be a solution of (6.21). Then $\hat{V}$ is a $(P,\mathbb{F})$-semimartingale and also satisfies (6.20).*

Proof. Recall that h and γ are bounded, so the $(P,\mathbb{F})$-martingale M, defined by

$$M_t = E_P\left[h(S_T,\eta_T) + \int_0^T \gamma(s,S_s,\eta_{s-},\hat{V}_{s-})\, ds \,\middle|\, \mathcal{F}_t \right], \quad t \in [0,T],$$

is bounded. Then $\hat{V}$ solves the (forward) stochastic differential equation

$$V_t = M_t - \int_0^t \gamma(s,S_s,\eta_{s-},V_{s-})\, ds,$$

and by Lemma 6.14 $\hat{V}$ is a $(P,\mathbb{F})$-semimartingale. In particular $\hat{V}$ is $\mathbb{F}$-adapted, and so it also satisfies (6.20). □

Corollary 6.15 and its proof show how one can express a backward stochastic differential equation of the type (6.20) and (6.21) in terms of a forward stochastic differential equation once a candidate solution is given.

Before we come to the first convergence result for V^n, we need some more preparation. Recall that by Lemma 6.11 filtrations generated by processes with independent increments converge in the sense of Definition 6.10. However in our setting there are two problems which arise. First the processes W^n and N^n, which generate the filtrations, have independent increments only under P'^n, and second, for every $n \in \mathbb{N}$ the processes involved are defined on different probability spaces, so that it makes no sense yet to speak of convergence of filtrations. The second point can be circumvented by changing to *one* fixed probability space with the help of the Skorokhod embedding theorem, whereas for the first point we analyze convergence under P'^n first. To that end recall the density processes Z of $P \sim P'$ and Z^n of $P^n \sim P'^n$ from (5.9) and (5.18).

For the discrete-time process V^n, given as the unique solution of (6.8), we get for $t \in [0,T]$

$$\begin{aligned} V_t^n &= E_{P^n}\left[h(S_T^n, \eta_T^n) + \int_0^T \gamma^n(s, S_{s-}^n, \eta_{s-}^n, V_{s-}^n)\, dL_s^n \middle| \mathcal{F}_t^n \right] - \int_0^t \gamma^n(s, S_{s-}^n, \eta_{s-}^n, V_{s-}^n)\, dL_s^n \\ &= \frac{1}{Z_t^n} E_{P'^n}\left[\left(h(S_T^n, \eta_T^n) + \int_0^T \gamma^n(s, S_{s-}^n, \eta_{s-}^n, V_{s-}^n)\, dL_s^n \right) Z_T^n \middle| \mathcal{F}_t^n \right] \\ &\quad - \int_0^t \gamma^n(s, S_{s-}^n, \eta_{s-}^n, V_{s-}^n)\, dL_s^n, \end{aligned}$$

by first using the fact that $\bar{\lambda}^{n\ell j}\, dL^n$ is the compensator of $dN^{n\ell j}$ by Lemma 5.8, and then using the Bayes formula for conditional expectations.

We define the processes Γ^n and Γ by

$$\Gamma_t^n = \int_0^t \gamma^n(s, S_{s-}^n, \eta_{s-}^n, V_{s-}^n)\, dL_s^n, \quad \Gamma_t = \int_0^t \gamma(s, S_{s-}, \eta_{s-}, V_{s-})\, ds, \quad t \in [0,T],$$

the random variables u^n and u by

$$u^n = h(S_T^n, \eta_T^n) + \Gamma_T^n, \quad u = h(S_T, \eta_T) + \Gamma_T,$$

and finally the processes Υ^n and Υ by

$$\Upsilon_t^n = E_{P'^n}[u^n Z_T^n | \mathcal{F}_t^n], \quad \Upsilon_t = E_{P'}[u Z_T | \mathcal{F}_t], \quad t \in [0,T],$$

so that V^n and V satisfy

$$V_t^n = \frac{\Upsilon_t^n}{Z_t^n} - \Gamma_t^n \quad \text{and} \quad V_t = \frac{\Upsilon_t}{Z_t} - \Gamma_t, \quad t \in [0,T]. \tag{6.23}$$

Theorem 6.16 *Under the assumptions of Theorem 6.6 we have*

$$\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, V^n, Z^n | P'^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta, V, Z | P').$$

Proof. For simplicity we write $X^n = (L^n, W^n, N^n, S^n, \eta^n)$ and $X = (L, W, N, S, \eta)$. The proof is divided into several steps.

1) Due to the convergence of $(L^n, W^n, N^n, S^n, \eta^n, Z^n)$ (see Proposition 5.10) and the tightness of $(V^n)_{n\in\mathbb{N}}$ we have by Lemma 1.42 that the sequence $\mathcal{L}(X^n, V^n, Z^n|P'^n)_{n\in\mathbb{N}}$ is tight. To study its cluster points, we fix a subsequence, which we denote for simplicity also by (X^n, V^n, Z^n), and assume that for a further subsubsequence

$$\mathcal{L}(X^n, V^n, Z^n|P'^n) \xrightarrow{w} \mathcal{L}(\bar{X}, \bar{V}, \bar{Z}|\bar{P}') \tag{6.24}$$

for some càdlàg processes $\bar{X} = (\bar{L}, \bar{W}, \bar{N}, \bar{S}, \bar{\eta})$, $\bar{V}, \bar{Z}$ on some probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{P}')$, and such that $\mathcal{L}(\bar{X}, \bar{Z}|\bar{P}') = \mathcal{L}(X, Z|P')$.

2) If we define $\bar{\Gamma}_t := \int_0^t \gamma(s, \bar{S}_s, \bar{\eta}_{s-}, \bar{V}_{s-})\, ds$ for $t \in [0,T]$ and $\bar{u} := h(\bar{S}_T, \bar{\eta}_T) + \bar{\Gamma}_T$, then along the above chosen subsubsequence we have joint convergence

$$\mathcal{L}(X^n, V^n, \Gamma^n, Z^n, u^n, Z^n_T|P'^n) \xrightarrow{w} \mathcal{L}(\bar{X}, \bar{V}, \bar{\Gamma}, \bar{Z}, \bar{u}, \bar{Z}_T|\bar{P}'), \tag{6.25}$$

with $\mathcal{L}(\bar{X}, \bar{Z}, \bar{Z}_T|\bar{P}') = \mathcal{L}(X, Z, Z_T|P')$. Indeed by the construction of Γ^n via γ^n and (6.10) we have

$$\begin{aligned}
\Gamma^n_t &= \int_0^t \gamma(s, S^n_{s-}, \eta^n_{s-}, V^n_{s-})\, dL^n_s + \frac{T}{n} \int_0^t \sum_{i,j=1}^m f^{ij}(s, S^n_{s-}, V^n_{s-})(1 - \lambda^{ij}(s, S^n_{s-}))\, dL^n_s \\
&=: \int_0^t H^n_{s-}\, dL^n_s + R^n_t.
\end{aligned}$$

Furthermore (6.24) and the continuity of γ imply

$$\mathcal{L}(X^n, V^n, H^n, Z^n|P'^n) \xrightarrow{w} \mathcal{L}(\bar{X}, \bar{V}, \bar{H}, \bar{Z}|\bar{P}'),$$

where $\bar{H}_t := \gamma(t, \bar{S}_t, \bar{\eta}_t, \bar{V}_t)$, $t \in [0,T]$, by Proposition B.4 and the continuous mapping theorem. Since $(L^n)_{n\in\mathbb{N}}$ is good, the definition of goodness (cf. Definition 1.48) implies

$$\mathcal{L}\left(X^n, V^n, \int H^n_-\, dL^n, Z^n|P'^n\right) \xrightarrow{w} \mathcal{L}\left(\bar{X}, \bar{V}, \int \bar{H}_-\, dL, \bar{Z}|\bar{P}'\right).$$

Finally since $\bar{\Gamma} = \int \bar{H}_- dL$ and since R^n satisfies the assumptions of Lemma 1.47 because f^{ij} and λ^{ij} are uniformly bounded, we have

$$\mathcal{L}(X^n, V^n, \Gamma^n, Z^n|P'^n) \xrightarrow{w} \mathcal{L}(\bar{X}, \bar{V}, \bar{\Gamma}, \bar{Z}|\bar{P}').$$

The joint convergence in (6.25) follows from the continuous mapping theorem. In fact, $(\bar{S}, \bar{\eta}, \bar{\Gamma}, \bar{Z})$ is $\bar{P}'$-stochastically continuous and so the set of discontinuities of the projection $\alpha \mapsto \alpha_T$ on $\mathbb{D}$ is a null set for the $\bar{P}'$-distribution of $(\bar{X}, \bar{V}, \bar{\Gamma}, \bar{Z})$. Finally h is continuous by assumption, so (6.25) follows by Proposition B.4.

3) We now apply the Skorokhod embedding theorem (cf. Theorem 1.39) in order to get almost sure convergence. There exists a probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}')$ which supports càdlàg processes

$\hat{X}^n = (\hat{L}^n, \hat{W}^n, \hat{N}^n, \hat{S}^n, \hat{\eta}^n)$, $\hat{V}^n, \hat{\Gamma}^n, \hat{Z}^n$ and $\hat{X} = (\hat{L}, \hat{W}, \hat{N}, \hat{S}, \hat{\eta})$, $\hat{V}, \hat{\Gamma}, \hat{Z}$ and random variables $\hat{u}^n$, $\hat{Z}^n_T$ and $\hat{u}$, $\hat{Z}_T$, such that

$$\mathcal{L}(\hat{X}^n, \hat{V}^n, \hat{\Gamma}^n, \hat{Z}^n, \hat{u}^n, \hat{Z}^n_T | \hat{P}') = \mathcal{L}(X^n, V^n, \Gamma^n, Z^n, u^n, Z^n_T | P'^n)$$

and

$$\mathcal{L}(\hat{X}, \hat{V}, \hat{\Gamma}, \hat{Z}, \hat{u}, \hat{Z}_T | \hat{P}') = \mathcal{L}(\bar{X}, \bar{V}, \bar{\Gamma}, \bar{Z}, \bar{u}, \bar{Z}_T | \bar{P}'),$$

and such that

$$(\hat{X}^n, \hat{V}^n, \hat{\Gamma}^n, \hat{Z}^n, \hat{u}^n, \hat{Z}^n_T) \to (\hat{X}, \hat{V}, \hat{\Gamma}, \hat{Z}, \hat{u}, \hat{Z}_T) \quad \hat{P}'\text{-a.s.}$$

for the product topology of the Skorokhod topology on the path space of $(\hat{X}, \hat{V}, \hat{\Gamma}, \hat{Z})$ and the usual topology on $\mathbb{R}^2$. Note that by Lemma 1.40 we have

$$\hat{\Gamma}_t = \int_0^t \gamma(s, \hat{S}_s, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds \quad \text{and} \quad \hat{u} = h(\hat{S}_T, \hat{\eta}_T) + \hat{\Gamma}_T. \tag{6.26}$$

4) In the approximating models as well as in the limit model the processes V^n and V depend crucially on the chosen filtration. Recall that in the n-th approximating model $(\Omega^n, \mathcal{F}^n, P'^n)$ is endowed with $\mathbb{F}^n = \mathbb{F}^{(W^n, N^n)}(P'^n)$, the P'^n-completion of the (right-continuous) filtration generated by (W^n, N^n), while $\mathbb{F} = \mathbb{F}^{(W,N)}(P')$ is the filtration on $(\Omega, \mathcal{F}, P')$.

We thus endow $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P}')$ with the filtrations $\hat{\mathbb{F}}^n := \mathbb{F}^{(\hat{W}^n, \hat{N}^n)}$ and $\hat{\mathbb{F}} := \mathbb{F}^{(\hat{W}, \hat{N})}$ (note that in contrast to the above $\hat{\mathbb{F}}^n$ and $\hat{\mathbb{F}}$ are not completed), and we have by Lemma 1.41

$$\hat{V}^n_t = \frac{1}{\hat{Z}^n_t} E_{\hat{P}'}[\hat{u}^n \hat{Z}^n_T | \hat{\mathcal{F}}^n_t] - \hat{\Gamma}^n_t =: \frac{1}{\hat{Z}^n_t} \hat{\Upsilon}^n_t - \hat{\Gamma}^n_t \quad \hat{P}'\text{-a.s.} \tag{6.27}$$

since $V^n_t = \frac{1}{Z^n_t} E_{P'}[u^n Z^n_T | \mathcal{F}^n_t] - \Gamma^n_t$ under P'^n. Now since $(\hat{W}^n, \hat{N}^n)$ are càdlàg processes with independent increments, and since $(\hat{W}^n, \hat{N}^n) \to (\hat{W}, \hat{N})$ $\hat{P}'$-a.s. and thus in probability, we have $\hat{\mathbb{F}}^n \xrightarrow{w} \hat{\mathbb{F}}$ by Lemma 6.11.

Now $E_{\hat{P}'}[\hat{Z}^n_T] = 1 = E_{\hat{P}'}[\hat{Z}_T]$ for all $n \in \mathbb{N}$, so nonnegativity and the $\hat{P}'$-a.s. convergence of $\hat{Z}^n_T$ to $\hat{Z}_T$ imply convergence in $\mathcal{L}^1(\hat{P}')$ (cf. Durrett (1991), Theorem 4.5.2), and since $\hat{u}^n, \hat{u}$ are bounded by the same constant, we have that $(\hat{Z}^n_T, \hat{u}^n \hat{Z}^n_T) \to (\hat{Z}_T, \hat{u}\hat{Z}_T)$ in $\mathcal{L}^1(\hat{P}')$. Therefore if we define the process $\hat{\Upsilon}$ by

$$\hat{\Upsilon}_t = E_{\hat{P}'}[\hat{u}\hat{Z}_T | \hat{\mathcal{F}}_t], \tag{6.28}$$

Lemma 6.12 yields

$$(\hat{Z}^n, \hat{\Upsilon}^n) \xrightarrow{\hat{P}'} (\hat{Z}, \hat{\Upsilon})$$

for the Skorokhod topology. Furthermore $\hat{\Gamma}^n \to \hat{\Gamma}$ $\hat{P}'$-a.s. and thus also in $\hat{P}'$-probability for the Skorokhod topology, and $\hat{\Gamma}$ is $\hat{P}'$-a.s. continuous. Therefore we have

$$(\hat{Z}^n, \hat{\Upsilon}^n, \hat{\Gamma}^n) \xrightarrow{\hat{P}'} (\hat{Z}, \hat{\Upsilon}, \hat{\Gamma})$$

for the Skorokhod topology. This in turn implies that there exists a (further) subsubsubsequence along which we have

$$(\hat{Z}^n, \hat{\Upsilon}^n, \hat{\Gamma}^n) \to (\hat{Z}, \hat{\Upsilon}, \hat{\Gamma}) \quad \hat{P}'\text{-a.s.} \tag{6.29}$$

for the Skorokhod topology, and since $(z, y, g) \mapsto \frac{y}{z} - g$ is continuous on $(0, \infty) \times \mathbb{R} \times \mathbb{R}$, we have with Proposition B.4

$$\frac{\hat{\Upsilon}^n}{\hat{Z}^n} - \hat{\Gamma}^n \to \frac{\hat{\Upsilon}}{\hat{Z}} - \hat{\Gamma} \quad \hat{P}'\text{-a.s.}$$

for the Skorokhod topology. On the other hand we have $\hat{V}^n \to \hat{V}$ $\hat{P}'$-a.s., so together with (6.27) we have

$$\hat{V} = \frac{\hat{\Upsilon}}{\hat{Z}} - \hat{\Gamma} \quad \hat{P}'\text{-a.s.},$$

which means by the definitions of $\hat{\Upsilon}$ and $\hat{\Gamma}$ in (6.28) and (6.26) that $\hat{V}$ satisfies

$$\begin{aligned} \hat{V}_t &= \frac{1}{\hat{Z}_t} E_{\hat{P}'} \left[\left(h(\hat{S}_T, \hat{\eta}_T) + \int_0^T \gamma(s, \hat{S}_{s-}, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds \right) \hat{Z}_T \middle| \hat{\mathcal{F}}_t \right] \\ &\quad - \int_0^t \gamma(s, \hat{S}_{s-}, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds, \quad t \in [0, T]. \end{aligned}$$

5) We now carry out an obvious change of measure in order to obtain an equation analogous to (6.21) on $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$, where $\hat{P} \ll \hat{P}'$ and where $\hat{S}$ and $\hat{\eta}$ admit the same mutual dependences under $\hat{P}$ as S and η under P. We have that $\mathcal{L}(\hat{N}, \hat{S}, \hat{Z} | \hat{P}') = \mathcal{L}(N, S, Z | P')$, so since

$$Z = \mathcal{E}\left(\sum_{\ell,j=1}^m \int \left(\lambda^{\ell j}(t, S_t) - 1 \right) \left(dN_t^{\ell j} - dt \right) \right)$$

by (5.9), we have by Lemma 1.40

$$\hat{Z} = \mathcal{E}\left(\sum_{\ell,j=1}^m \int \left(\lambda^{\ell j}(t, \hat{S}_t) - 1 \right) \left(d\hat{N}_t^{\ell j} - dt \right) \right). \tag{6.30}$$

Therefore $\hat{Z}$ is a strictly positive $(\hat{P}', \hat{\mathbb{F}})$-martingale, and if we define $d\hat{P} = \hat{Z}_T\, d\hat{P}'$, we get with the Bayes formula for conditional expectations that $\hat{V}$ satisfies

$$\hat{V}_t = E_{\hat{P}} \left[h(\hat{S}_T, \hat{\eta}_T) + \int_0^T \gamma(s, \hat{S}_{s-}, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds \middle| \hat{\mathcal{F}}_t \right] - \int_0^t \gamma(s, \hat{S}_{s-}, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds$$

for all $t \in [0, T]$. Note that (6.30) yields with the same arguments as in Section 5.1 that $\hat{N}_t^{\ell j}$ has $(\hat{P}', \hat{\mathbb{F}})$-intensities $\lambda^{\ell j}(t, \hat{S}_t)$ for $\ell, j \in \{1, \ldots, m\}$.

6) It remains to show that $\hat{V}$ solves an equation corresponding to (6.20) for $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{F}}, \hat{P})$. Since h and γ are bounded functions, $\hat{V}$ is bounded, and $\hat{V}$ has càdlàg paths by 3). Furthermore as

the a.s. limit of $\mathbb{F}^n$-adapted processes, $\hat{V}$ is adapted to $\hat{\mathbb{G}} = \bigvee_{n\in\mathbb{N}} \hat{\mathbb{F}}^n$ and thus $\hat{\mathcal{F}} \otimes \mathcal{B}([0,T])$-measurable, so $\hat{V}$ is in $\mathcal{X}_b$, and Corollary 6.15 yields that $\hat{V}$ satisfies

$$\hat{V}_t = E_{\hat{P}}\left[h(\hat{S}_T, \hat{\eta}_T) + \int_t^T \gamma(s, \hat{S}_{s-}, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds \middle| \hat{\mathcal{F}}_t \right].$$

Notice that this is not quite yet the analogue of (6.20), since in the original continuous-time model on $(\Omega, \mathcal{F}, P)$ we have assumed that the filtration satisfies the usual conditions, whereas $\hat{\mathbb{F}}$ is merely right-continuous. However by Lemma C.10 for the conditional expectation it is irrelevant whether the σ-algebra involved is complete or not, so that $\hat{V}$ also satisfies

$$\hat{V}_t = E_{\hat{P}}\left[h(\hat{S}_T, \hat{\eta}_T) + \int_t^T \gamma(s, \hat{S}_{s-}, \hat{\eta}_{s-}, \hat{V}_{s-})\, ds \middle| \hat{\mathcal{F}}_t(\hat{P}) \right], \tag{6.31}$$

where by $\hat{\mathbb{F}}(\hat{P})$ we denote the $\hat{P}$-augmentation of $\hat{\mathbb{F}}$. In particular, $\hat{V}$ is adapted to $\hat{\mathbb{F}}$ and $\hat{\mathbb{F}}(\hat{P})$.

7) We are finally in a position to prove the convergence claimed in Theorem 6.16. The preceding arguments show that every subsequence of $\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, V^n, Z^n | P'^n)$ contains a further subsequence which converges weakly to $\mathcal{L}(\hat{L}, \hat{W}, \hat{N}, \hat{S}, \hat{\eta}, \hat{V}, \hat{Z} | \hat{P}')$, where $\hat{V}$ satisfies (6.31) and where

$$\mathcal{L}(\hat{L}, \hat{W}, \hat{N}, \hat{S}, \hat{\eta}, \hat{Z} | \hat{P}') = \mathcal{L}(L, W, N, S, \eta, Z | P').$$

Now by Remark 6.1 and because $\hat{V}$ is $\hat{\mathbb{F}}(\hat{P})$-adapted and a bounded $\hat{\mathbb{F}}(\hat{P})$-semimartingale, we have $\hat{V}_t = v(t, \hat{S}_t, \hat{\eta}_t)$ and $V_t = v(t, S_t, \eta_t)$, where the function v as the solution of the reaction-diffusion equation (6.5) is independent of the choice of the underlying probability space. This yields that for the measurable functional $w\colon \mathbb{D}((0,\infty)^d) \times \mathbb{D}(\{1,\ldots,m\}) \to \mathbb{D}((0,\infty))$, $w(\alpha,\beta)(t) = v(t, \alpha(t), \beta(t))$, we have that $V = w(S,\eta)$ and $\hat{V} = w(\hat{S}, \hat{\eta})$. So we immediately get from $\mathcal{L}(\hat{X}, \hat{Z} | \hat{P}') = \mathcal{L}(X, Z | P')$ that

$$\mathcal{L}(\hat{L}, \hat{W}, \hat{N}, \hat{S}, \hat{\eta}, \hat{V}, \hat{Z} | \hat{P}') = \mathcal{L}(L, W, N, S, \eta, V, Z | P').$$

Hence every cluster point has the same distribution, and so the asserted convergence follows. □

Corollary 6.17 *Under the assumptions of Theorem 6.6 we have*

$$\mathcal{L}(L^n, W^n, N^n, S^n, \eta^n, V^n, Z^n | P^n) \xrightarrow{w} \mathcal{L}(L, W, N, S, \eta, V, Z | P).$$

Proof. By Theorem 6.16 we have convergence under P'^n and P', and $(P^n)_{n\in\mathbb{N}} \overset{\text{loc}}{\lhd} (P'^n)_{n\in\mathbb{N}}$ by Proposition 5.11, so Theorem 1.56 yields the result. □

Remark 6.18 We have seen that for bounded and sufficiently smooth payoff functions the price process of the payoff structure in the continuous-time model may be approximated by the price processes in the discrete-time models. Another possibility to show convergence of the price processes is to show convergence of the functions v^n, given by the backward computation scheme from Section 6.2, to the function v, given as the solution of the reaction-diffusion equation (6.5), in a suitable sense. But by the structure of the functions v^n it seems to be delicate to show their convergence to v. However such a convergence result would imply convergence of certain hedging strategies. Becherer and Schweizer (2003) show that the payoff structure B may be decomposed as

$$B = E_P[B] + \int_0^T \vartheta_s \, dS_s + L_T,$$

where L is a sum of stochastic integrals with respect to the P-martingales $N_t^{\ell j} - \int_0^t \lambda^{\ell j}(s, S_s)\, ds$, $\ell, j \in \{1, \ldots, m\}$, and that the locally risk-minimizing strategy ϑ and the hedge error L for B is given explicitly by v and its derivatives. Now in the n-th discrete-time model the Kunita-Watanabe decomposition yields that the payoff structure B^n from 6.9 can be written as

$$B^n = E_{P^n}[B^n] + \int_0^T \vartheta_s^n \, dS_s^n + L_T^n,$$

where L^n is P^n-orthogonal to S^n. Furthermore the predictable process ϑ^n is the locally risk-minimizing strategy for B^n, since P^n is a martingale measure for S^n. So if we show an analogous dependence of ϑ_s^n and L^n on v^n and the derivatives of v^n, then $v^n \to v$ in a suitable sense implies convergence of the locally risk-minimizing strategies and hedge errors. This is left as a topic for further research. $\diamond$

APPENDIX

Appendix A

Infinitely Divisible Distributions and Lévy Processes

Here we recall properties of infinitely divisible distributions and their relation with Lévy processes. These properties allow us to construct absolutely continuous measures from given Girsanov quantities in a different way than we have seen in Corollary 2.17. This is shown in Theorem A.9 at the end of this part of the appendix.

Definition A.1 A distribution μ on $\mathbb{R}^d$ is called *infinitely divisible* if for any $n \in \mathbb{N}$ there exists a distribution μ_n on $\mathbb{R}^d$ such that $\mu = \mu_n^{\otimes n}$. $\diamond$

Remark A.2 Definition A.1 is equivalent to the following characterizations of infinitely divisible distributions:

a) If μ is the distribution of an $\mathbb{R}^d$-valued random variable X, then μ is infinitely divisible if and only if for any $n \in \mathbb{N}$ there exists a family of i.i.d. $\mathbb{R}^d$-valued random variables $(X_k^n)_{k=1,\ldots,n}$ such that $X = \sum_{k=1}^n X_k^n$ P-a.s.

b) If φ is the characteristic function of μ (i.e. $\varphi_t(u) = \int e^{iu^{\mathrm{tr}}x}\ \mu(dx)$), then μ is infinitely divisible if and only if for any $n \in \mathbb{N}$ there exists a characteristic function φ_n such that $\varphi(u) = (\varphi_n(u))^n$ for all $u \in \mathbb{R}^d$. $\diamond$

Theorem A.3 *Let h be a fixed truncation function.*

a) Let μ be infinitely divisible. Then its characteristic function is of the form

$$\varphi(z) = \exp\left(ib^{\mathrm{tr}}z - \frac{1}{2}z^{\mathrm{tr}}cz + \int_{\mathbb{R}^d}\left(e^{iz^{\mathrm{tr}}x} - 1 - iz^{\mathrm{tr}}h(x)\right)\ K(dx)\right), \tag{A.1}$$

where $b = b(h) \in \mathbb{R}^d$, c is a nonnegative definite symmetric matrix, and K is a measure on $(\mathbb{R}^d, \mathcal{B}^d)$ satisfying $K(\{0\}) = 0$ and $\int_{\mathbb{R}^d}(|z|^2 \wedge 1)\ K(dz) < \infty$.

b) *Conversely, take* $b \in \mathbb{R}^d$*, a nonnegative definite symmetric matrix* c*, and a measure* K *on* $(\mathbb{R}^d, \mathcal{B}^d)$ *satisfying* $K(\{0\}) = 0$ *and* $\int_{\mathbb{R}^d}(|z|^2 \wedge 1)\ K(dz) < \infty$*. Then there exists an infinitely divisible distribution* μ *whose characteristic function is given by (A.1).*

Proof. cf. Sato (1999), Theorem 8.1. □

Definition A.4 The triplet (b, c, K) from Theorem A.3 of an infinitely divisible distribution μ is called *characteristic triplet* of μ. ◇

Now let L be a Lévy process. From the independence and stationarity of the increments of L it follows that L_t is infinitely divisible for every $t \geq 0$. In fact, for each n we can take the i.i.d. increments $L_k^n := L_{\frac{kt}{n}} - L_{\frac{(k-1)t}{n}}$, $k = 1, \ldots, n$, to obtain $L_t = \sum_{k=1}^n L_k^n$.

Let φ_t be the characteristic function of the distribution of L_t. Then it immediately follows from the above that

$$\varphi_{\frac{p}{q}}(z) = (\varphi_1(z))^{\frac{p}{q}}$$

for $p, q \in \mathbb{N}$. For $t \geq 0$, take a sequence $t_n \in \mathbb{Q}$ with $t_n \to t$. Then by the stochastic continuity of L it follows that $L_{t_n} \to L_t$ in probability which implies that the distribution of L_{t_n} converges weakly to the distribution of L_t, and thus $\varphi_{t_n} \to \varphi_t$ pointwise. This implies $\varphi_t(z) = (\varphi_1(z))^t$ for all $t \geq 0$. We have proved

Theorem A.5 *Let* L *be a Lévy process and* (b, c, K) *the characteristic triplet from Theorem A.3 of the distribution of* L_1*. Then the characteristic function of* L_t *is of the form*

$$\varphi_t(z) = \exp\left(t\Big(ib^{\mathrm{tr}}z - \frac{1}{2}z^{\mathrm{tr}}cz + \int_{\mathbb{R}^d}\left(e^{iz^{\mathrm{tr}}x} - 1 - iz^{\mathrm{tr}}h(x)\right)\ K(dx)\Big)\right) \tag{A.2}$$

for all $t \geq 0$*.*

Formula (A.1) in Theorem A.3 is also known as the Lévy-Khinchine formula. Note that (b, c, K) in Theorem A.5 are the Lévy characteristics of L as defined in Theorem 1.27. In order to construct absolutely continuous measures which preserve the Lévy property we need the following converse statement to Theorem A.5 which says that, loosely speaking, to any infinitely divisible distribution there exists a Lévy process with this distribution.

Theorem A.6 *Let* $(\Omega, \mathcal{F}) = (\mathbb{D}, \mathcal{B}(\mathbb{D}))$ *and let* L *be the coordinate process on* $\mathbb{D}$*. To any infinitely divisible distribution* μ *on* $\mathbb{R}^d$ *there exists a probability measure* P *on* $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ *under which* L *is a Lévy process with* $\mathcal{L}(L_1|P) = \mu$*.*

Proof. cf. Sato (1999), Corollary 11.6. □

Theorem A.5 implies that the characteristic function $\varphi_t(u) = E[\exp(iu^{\mathrm{tr}}L_t)]$ of L_t has the property

$$\varphi_t(u) = (\varphi_1(u))^t.$$

The following result shows that this also holds for the corresponding moment generating function $m_t(u) = E[\exp(u^{\mathrm{tr}} L_t)]$ as long as it exists.

Theorem A.7 *Let L be a Lévy process with Lévy characteristics (b, c, K). Let*

$$A = \left\{ u \in \mathbb{R}^d \colon \int_{\{|x|>1\}} e^{u^{\mathrm{tr}} x} \; K(dx) < \infty \right\}.$$

Then

a) The set A is convex and contains the origin.

b) $u \in A$ if and only if $E[\exp(u^{\mathrm{tr}} L_t)] < \infty$ for some $t > 0$.

c) If $u \in \mathbb{C}^d$ such that $\Re(u) \in A$ then

$$\Psi(u) = b^{\mathrm{tr}} u + \frac{1}{2} u^{\mathrm{tr}} c u + \int_{\mathbb{R}^d} \left(e^{u^{\mathrm{tr}} x} - 1 - (u^{\mathrm{tr}} x) \mathbb{1}_{\{|x| \le 1\}} \right) \; K(dx)$$

is well-defined in $\mathbb{C}$, $E\big[\big|e^{u^{\mathrm{tr}} L_t}\big|\big] < \infty$ and

$$E\big[e^{u^{\mathrm{tr}} L_t}\big] = e^{t \Psi(u)}$$

for all $t \in [0, \infty)$.

Proof. cf. Sato (1999), Theorem 25.17. □

Theorem A.7 implies that the characteristic function of the distribution of L_t can be analytically extended to some "horizontal" strip in $\mathbb{C}^d$ if $A \neq \{0\}$, and that the moment generating function of L_t coincides with this extension along the imaginary axis.

An Alternative Construction of Measures from Given Girsanov Quantities

In Corollary 2.17 we have constructed from *given* Girsanov quantities locally equivalent measures which preserve the Lévy structure of a given Lévy process. Here we give an alternative construction via infinitely divisible distributions. Theorem A.3 b) gives us criteria under which conditions given quantities (b, c, K) form the characteristic triplet of an infinitely divisible distribution, and from Theorem A.6 we know that to every infinitely divisible distribution there exists a corresponding Lévy process (i.e. a distribution on $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ under which the coordinate process is a Lévy process). The following theorem gives us conditions for two measures, under which a process L is a Lévy process, to be locally equivalent.

Theorem A.8 *Let $(\Omega, \mathcal{F}, \mathbb{F})$ be a filtered measurable space, L a stochastic process and P and P' two probability measures on $(\Omega, \mathcal{F})$. Suppose L is a Lévy process under both P and P' with Lévy characteristics (b, c, K) and (b', c', K') relative to a truncation function h, respectively. Then $P' \stackrel{\text{loc}}{\ll} P$ if and only if the following conditions hold:*

(i) $K'(dx) = k(x)\ K(dx)$ *for some Borel function* $k \colon \mathbb{R}^d \to [0, \infty)$*;*

(ii) $\int_{\mathbb{R}^d} |h(x)(k(x) - 1)|\ K(dx) < \infty$*;*

(iii) $b' = b + c\beta + \int_{\mathbb{R}^d} h(x)(k(x) - 1)\ K(dx)$ *for some* $\beta \in \mathbb{R}^d$*;*

(iv) $c' = c$*;*

(v) $\int_{\mathbb{R}^d} \left(1 - \sqrt{k(x)}\right)^2\ K(dx) < \infty$*.*

Proof. cf. Jacod and Shiryaev (1987), Theorem IV.4.39 c). □

We can now construct locally equivalent measures which preserve the Lévy property of a given Lévy process. Recall the convex functions $f, g \colon [0, \infty) \to \mathbb{R}$ which are defined by

$$\begin{aligned} f(y) &= \begin{cases} y \log y - (y - 1) & \text{if } y > 0 \\ 1 & \text{if } y = 0, \end{cases} \\ g(y) &= (1 - \sqrt{y})^2, \end{aligned}$$

and recall that $0 \le g(y) \le f(y)$ for all $y \ge 0$ by Lemma 2.13.

Theorem A.9 *Let L be the coordinate process on $\mathbb{D}$ and let P be a probability measure on $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ such that L is a P-Lévy process with P-Lévy characteristics (b, c, K) relative to some truncation function h. Fix $\bar{\beta} \in \mathbb{R}^d$ and let $\bar{Y} \colon \mathbb{R}^d \to (0, \infty)$ be measurable with*

$$\int_{\mathbb{R}^d} f(\bar{Y}(x))\ K(dx) < \infty. \tag{A.3}$$

Then $\int_{\mathbb{R}^d} |h(x)(\bar{Y}(x) - 1)|\ K(dx) < \infty$, and if we define

$$\begin{aligned} \bar{b} &= b + c\bar{\beta} + \int_{\mathbb{R}^d} h(x) \left(\bar{Y}(x) - 1\right)\ K(dx), \\ \bar{c} &= c, \\ \bar{K}(dx) &= \bar{Y}(x)\ K(dx), \end{aligned}$$

there exists a probability measure $\bar{Q} \stackrel{\text{loc}}{\sim} P$ on $(\mathbb{D}, \mathcal{B}(\mathbb{D}), \mathbb{F}^{\mathbb{D}})$, where $\mathbb{F}^{\mathbb{D}}$ is the P-completion of the canonical filtration on $\mathbb{D}$, under which L is a Lévy process with Lévy characteristics $(\bar{b}, \bar{c}, \bar{K})$.

Proof. We want to use Theorem A.3 b) in order to obtain an infinitely divisible distribution on $\mathbb{R}^d$ with characteristic triplet $(\bar{b}, \bar{c}, \bar{K})$ and then construct $\bar{Q} \overset{\text{loc}}{\sim} P$ via Theorems A.6 and A.8. From Lemma C.5 we know that for $\nu(ds, dx) = dsK(dx)$ and some $t > 0$

$$\int_{\mathbb{R}^d} |h(x)(\bar{Y}(x) - 1)| \ K(dx) = \frac{1}{t}|h \ (\bar{Y} - 1)| * \nu_t \leq \text{const.} \left(1 + \int_{\mathbb{R}^d} f(\bar{Y}(x)) \ K(dx)\right)$$

which is finite by assumption, so that $\bar{b} \in \mathbb{R}^d$ is well-defined. Since c is a nonnegative definite matrix, so is $\bar{c}$, and it remains to show that the assumptions on $\bar{K}$ of Theorem A.3 b) hold. First, $\bar{K}$ is a measure on $\mathbb{R}^d$, since $\bar{Y} > 0$. Furthermore $\bar{Y}(0) < \infty$ and $K(\{0\}) = 0$ by assumption, so $\bar{K}(\{0\}) = \int_{\{0\}} \bar{Y}(x) \ K(dx) = 0$. So it remains to show the integrability condition $\int_{\mathbb{R}^d} (1 \wedge |x|^2) \ \bar{K}(dx) = \int_{\mathbb{R}^d} (1 \wedge |x|^2) \bar{Y}(x) \ K(dx) < \infty$. The measure $\widetilde{K}$, defined by $\widetilde{K}(dx) = (1 \wedge |x|^2) \ K(dx)$, is a finite measure by assumption, so for $C = \widetilde{K}(\mathbb{R}^d) > 0$, $\frac{1}{C}\widetilde{K}$ is a probability measure. (If $C = 0$, then $K \equiv 0$ and there is nothing to show.) Then, since f is convex, we get by Jensen's inequality and the fact that $\frac{d\widetilde{K}}{dK} \leq 1$

$$f\left(\frac{1}{C}\int_{\mathbb{R}^d} \bar{Y}(x) \ \widetilde{K}(dx)\right) \leq \frac{1}{C}\int_{\mathbb{R}^d} f(\bar{Y}(x)) \ \widetilde{K}(dx) \leq \frac{1}{C}\int f(\bar{Y}(x)) \ K(dx) < \infty$$

by (A.3). Now $f(y) < \infty$ implies $y < \infty$, so that

$$\int_{\mathbb{R}^d} (1 \wedge |x|^2) \ \bar{K}(dx) = \int_{\mathbb{R}^d} \bar{Y}(x) \widetilde{K}(dx) < \infty$$

since $C > 0$.

We have shown that the given triplet $(\bar{b}, \bar{c}, \bar{K})$ defines an infinitely divisible distribution on $\mathbb{R}^d$, so by Theorem A.6 there exists a probability measure $\bar{Q}$ on $(\mathbb{D}, \mathcal{B}(\mathbb{D}))$ under which the coordinate process L is a Lévy process with Lévy characteristics $(\bar{b}, \bar{c}, \bar{K})$. It remains to show that the conditions in Theorem A.8 hold (in both directions), so that $\bar{Q} \overset{\text{loc}}{\sim} P$. We start with $\bar{Q} \overset{\text{loc}}{\ll} P$. Note that above we have already shown (ii) (with $k = \bar{Y}$), whereas (i), (iii) and (iv) hold by construction (with $\beta = \bar{\beta}$). (v) is immediate by (A.3) and the fact that $\left(1 - \sqrt{\bar{Y}}\right)^2 = g(\bar{Y}) \leq f(\bar{Y})$ by Lemma 2.13. In order to show $P \overset{\text{loc}}{\ll} \bar{Q}$, take $k(x) = \frac{1}{\bar{Y}(x)}$ and $\beta = -\bar{\beta}$; then it is immediate that (i)–(v) hold. $\square$

Appendix B

The Skorokhod Topology

The results of Part III heavily depend on the continuous mapping theorem applied to sequences of random variables with values in the Skorokhod space. So we recall some results on the Skorokhod topology and show the Skorokhod-continuity of certain mappings on the Skorokhod space.

Let (E, d_E) be a Polish space, and denote by $\mathbb{D}(E)$ (or simply $\mathbb{D}$ if there is no ambiguity about E) the space of all càdlàg functions $\alpha\colon \mathbb{R}_+ \to E$. For simplicity (and since we only consider Polish spaces E which are subspaces of $\mathbb{R}^d$, endowed with the metric $d_0(x,y) = |x-y|$), we denote the metric on E by $d_E(x,y) := |x-y|$.

Let $\Lambda = \{\lambda\colon \mathbb{R}_+ \to \mathbb{R}_+ \text{ strictly increasing}, \lambda(0) = 0, \lim_{t\to\infty} \lambda(t) = \infty\}$. Recall that the Skorokhod topology on $\mathbb{D}$ is a metrizable topology such that a sequence $(\alpha_n)_{n\in\mathbb{N}}$ converges to α with respect to this topology if and only if there exists a sequence $(\lambda_n)_{n\in\mathbb{N}} \subseteq \Lambda$ such that

$$\text{(B.1)} \qquad \begin{cases} \text{(i)} & \sup\limits_{t\geq 0} |\lambda_n(t) - t| \overset{n\to\infty}{\longrightarrow} 0, \\ \text{(ii)} & \sup\limits_{t\leq N} |\alpha_n \circ \lambda_n(t) - \alpha(t)| \overset{n\to\infty}{\longrightarrow} 0 \text{ for all } N \in \mathbb{N}. \end{cases}$$

In the sequel we denote convergence with respect to the Skorokhod topology by $\alpha_n \xrightarrow{S} \alpha$ and we call $\lambda \in \Lambda$ a *time change*. See Jacod and Shiryaev (1987), Sections VI.1 and VI.2 for a detailed description of the Skorokhod topology and a number of continuous functions on $\mathbb{D}$. In the sequel we expand this list of continuous functions on $\mathbb{D}$, and we start with the following preparatory lemmas.

Lemma B.1 *Let $E \subseteq \mathbb{R}^d$, $E' \subseteq \mathbb{R}^{d'}$ be Polish spaces and let $\alpha_n \xrightarrow{S} \alpha$ in $\mathbb{D}(E)$ and $\beta_n \xrightarrow{S} \beta$ in $\mathbb{D}(E')$ for the same sequence of time changes. Then $(\alpha_n, \beta_n) \xrightarrow{S} (\alpha, \beta)$ in $\mathbb{D}(E \times E')$.*

Proof. This is immediate from the characterization of convergence in the Skorokhod topology, in particular (B.1 ii). □

Lemma B.2 *Let $\alpha_n \xrightarrow{S} \alpha$ in $\mathbb{D}(\mathbb{R}^d)$ and $\beta_n \xrightarrow{S} \beta$ in $\mathbb{D}(E')$ for some Polish space $E' \subseteq \mathbb{R}^{d'}$. Let $(\lambda_n^\beta)_{n\in\mathbb{N}}$ satisfy (B.1 ii) for β_n, β. If α is continuous, then $(\lambda_n^\beta)_{n\in\mathbb{N}}$ meets (B.1 ii) for α_n, α as well, and we have $(\alpha_n, \beta_n) \xrightarrow{S} (\alpha, \beta)$ in $\mathbb{D}(\mathbb{R}^d \times E')$.*

Proof. This follows from Jacod and Shiryaev (1987), Propositions VI.1.23, VI.2.2 b) and VI.2.1 a) and the proof thereof. □

Corollary B.3 *Suppose φ is continuous on $\mathbb{D}(E')$ for some Polish space $E' \subseteq \mathbb{R}^{d'}$. Then the mapping Φ on $\mathbb{D}(\mathbb{R}^d \times E')$, defined by $\Phi(\alpha, \beta) = (\alpha, \varphi(\beta))$ is continuous in all points $(\alpha, \beta) \in \mathbb{D}(\mathbb{R}^d \times E')$ such that α is continuous.*

Proof. This is an immediate consequence of Lemma B.1 and B.2. □

The following result seems to be folklore, but we have not been able to find it in the literature.

Proposition B.4 *Let $\varphi\colon \mathbb{R}^d \to \mathbb{R}^{d'}$ be continuous. Then $\Phi\colon \mathbb{D}(\mathbb{R}^d \times E) \to \mathbb{D}(\mathbb{R}^d \times E \times \mathbb{R}^{d'})$, defined by $(\Phi(\alpha, \beta))(t) = (\alpha(t), \beta(t), \varphi(\alpha(t)))$, is continuous on $\mathbb{D}(\mathbb{R}^d \times E)$.*

Proof. We show that $(\alpha_n, \beta_n) \xrightarrow{S} (\alpha, \beta)$ implies $\Phi(\alpha_n, \beta_n) \xrightarrow{S} \Phi(\alpha, \beta)$. To that end let $N \in \mathbb{N}$, $\varepsilon > 0$ and suppose $(\alpha_n, \beta_n) \xrightarrow{S} (\alpha, \beta)$ with time change $(\lambda_n)_{n\in\mathbb{N}}$. Since α has only countably many jumps, α is bounded on the interval $[0, N]$, and we set $m_N = \sup_{s\le N} |\alpha(s)|$ and $K_N = \{x \in \mathbb{R}^d \mid |x| \le 2m_N\}$. Then φ is uniformly continuous on K_N, so there exists $\delta > 0$ such that $|x - y| < \delta$ implies $|\varphi(x) - \varphi(y)| < \varepsilon$ for all $x, y \in K_N$.

Now for n_1 sufficiently large we have for all $n \ge n_1$

$$\sup_{s\le N} |\alpha_n \circ \lambda_n(s)| \le \sup_{s\le N} |\alpha_n \circ \lambda_n(s) - \alpha(s)| + \sup_{s\le N} |\alpha(s)| \le 2m_N,$$

since $\alpha_n \xrightarrow{S} \alpha$, and thus $|\alpha_n \circ \lambda_n(s) - \alpha(s)| \to 0$ as $n \to \infty$. So for $n \ge n_1$ we have $\alpha_n \circ \lambda_n(s), \alpha(s) \in K_N$ for all $s \le N$.

Furthermore $\alpha_n \xrightarrow{S} \alpha$ implies that for the above δ there exists n_2 such that for all $n \ge n_2$ we have $|\alpha_n \circ \lambda_n(s) - \alpha(s)| < \delta$ for all $s \le N$, so that the uniform continuity of φ on K_N implies that for all $n \ge n_0 = \max\{n_1, n_2\}$ we have $|\varphi(\alpha_n \circ \lambda_n(s)) - \varphi(\alpha(s))| < \varepsilon$ for all $s \le N$, and thus $\sup_{s\le N} |\varphi(\alpha_n \circ \lambda_n(s)) - \varphi(\alpha(s))| < \varepsilon$. Hence $\varphi(\alpha_n) \xrightarrow{S} \varphi(\alpha)$ with time change $(\lambda_n)_{n\in\mathbb{N}}$, so Lemma B.1 implies that $\Phi(\alpha_n, \beta_n) = (\alpha_n, \beta_n, \varphi(\alpha_n)) \xrightarrow{S} (\alpha, \beta, \varphi(\alpha)) = \Phi(\alpha)$. □

The following propositions show continuity of some functions on the Skorokhod space which involve “stopping times” (note that we do not consider a filtration on $\mathbb{D}$, so we put “stopping times” in quotation marks).

For any finite subset J of $\mathbb{R}^d$ let $\delta_J = \min\{|\ell - j|\colon \ell, j \in J\}$ and define for $\alpha \in \mathbb{D}(J)$ the jump times by

$$\text{(B.2)} \qquad \begin{cases} \tau_0(\alpha) &= 0, \\ \tau_k(\alpha) &= \inf\left\{t > \tau_{k-1} \mid |\Delta\alpha_t| > \frac{\delta_J}{2}\right\}. \end{cases}$$

Note that α takes values only in the discrete set J, so the jump sizes are at least of size $\delta_J > 0$, hence $(\tau_k(\alpha))_{k\in\mathbb{N}}$ indeed captures all jumps of α.

For Propositions B.5–B.7 let $E' \subseteq \mathbb{R}^{d'}$ and $E'' \subseteq \mathbb{R}^{d''}$ be Polish spaces for some $d', d'' \in \mathbb{N}$ and let I be a finite subset of $\mathbb{R}$; think of $I = \{\pi\} \cup \{1, \ldots, m\}$ from Part III. Note that convergence $(\alpha_n, \beta_n, \gamma_n) \xrightarrow{S} (\alpha, \beta, \gamma)$ in $\mathbb{D}(I \times E' \times E'')$ always means that both (α, β, γ) *and* $(\alpha_n, \beta_n, \gamma_n) \in \mathbb{D}(I \times E' \times E'')$ for all $n \in \mathbb{N}$.

Proposition B.5 *For all $k \in \mathbb{N}$ the functions $\varphi_k\colon \mathbb{D}(I \times E' \times E'') \to \mathbb{R}_+$, defined by $\varphi_k(\alpha, \beta, \gamma) = \tau_k(\alpha)$, are continuous on $\mathbb{D}(I \times E' \times E'')$.*

Proof. We show that $(\alpha_n, \beta_n, \gamma_n) \xrightarrow{S} (\alpha, \beta, \gamma)$ in $\mathbb{D}(I \times E' \times E'')$ implies $\tau_k(\alpha_n) \to \tau_k(\alpha)$. Now $(\alpha_n, \beta_n, \gamma_n) \xrightarrow{S} (\alpha, \beta, \gamma)$ implies $\alpha_n \xrightarrow{S} \alpha$ by projection, and the mapping $\alpha \mapsto \tau_k(\alpha)$ is continuous (cf. Jacod and Shiryaev (1987), Proposition VI.2.7; note that $\tau_k(\alpha) = t_k(\alpha, \frac{1}{2})$ and $\frac{1}{2} \notin U(\alpha)$ in the notation there). □

Proposition B.6 *For all $k \in \mathbb{N}$ the function $\varphi_k\colon \mathbb{D}(I \times E' \times E'') \to \mathbb{D}(I \times E' \times E'')$, defined by*

$$\varphi_k(\alpha, \beta, \gamma)(t) = (\alpha(t), \beta(t), \gamma(t \wedge \tau_k(\alpha)))$$

is continuous in every (α, β, γ) such that γ is continuous.

Proof. Let $(\alpha^n, \beta^n, \gamma^n) \xrightarrow{S} (\alpha, \beta, \gamma)$ for γ continuous, and let $\tau_k^n = \tau_k(\alpha_n)$ and $\tau_k = \tau_k(\alpha)$. Obviously the stopped function $\gamma^{\tau_k} = \gamma(\cdot \wedge \tau_k)$ is continuous since γ is continuous, so if $(\gamma^n)^{\tau_k^n} \to \gamma^{\tau_k}$ locally uniformly, then $(\gamma^n)^{\tau_k^n} \xrightarrow{S} \gamma^{\tau_k}$, and the claim follows from Lemma B.2. For fixed $N \in \mathbb{N}$ we have

$$\begin{aligned} \sup_{t\le N} |\gamma_n(t \wedge \tau_k^n) - \gamma(t \wedge \tau_k)| &\le \sup_{t\le N} |\gamma_n(t \wedge \tau_k^n) - \gamma(t \wedge \tau_k^n)| + \sup_{t\le N} |\gamma(t \wedge \tau_k^n) - \gamma(t \wedge \tau_k)| \\ &\le \sup_{t\le N} |(\gamma_n - \gamma)(t)| + \sup_{t\le N} |\gamma(t \wedge \tau_k^n) - \gamma(t \wedge \tau_k)|, \end{aligned}$$

which converges to 0 by Proposition B.5, since γ^{τ_k} is continuous and thus uniformly continuous on $[0, N]$. □

The next proposition is concerned with "choosing" a càdlàg function γ^ℓ among several functions $(\gamma^j)_{j\in I}$ depending on the value of a further function α at a certain "stopping time". If γ takes values in $(\mathbb{R}^d)^{|I|}$ we write $\gamma = (\gamma^j)_{j\in I}$, where each γ^j takes values in $\mathbb{R}^d$.

Proposition B.7 *The function* $\varphi_k\colon \mathbb{D}(I \times E' \times (\mathbb{R}^d)^{|I|}) \to \mathbb{D}(I \times E' \times \mathbb{R}^d)$*, given by*

$$\varphi_k(\alpha,\beta,\gamma)(t) = \left(\alpha(t), \beta(t), \sum_{\ell \in I} \gamma^\ell(t) \mathbb{1}_{\{\alpha(\tau_k(\alpha))=\ell\}}\right)$$

is continuous for the Skorokhod topology on $\mathbb{D}(I \times E' \times (\mathbb{R}^d)^{|I|})$ *for all* $k \in \mathbb{N}$*.*

Proof. We have to show that $(\alpha_n, \beta_n, \gamma_n) \xrightarrow{S} (\alpha,\beta,\gamma)$ in $\mathbb{D}(I \times E' \times E'')$ implies that $\varphi_k(\alpha_n,\beta_n,\gamma_n) \xrightarrow{S} \varphi_k(\alpha,\beta,\gamma)$. Now $(\alpha_n,\beta_n,\gamma_n) \xrightarrow{S} (\alpha,\beta,\gamma)$ if and only if there exists a sequence $(\lambda_n)_{n\in\mathbb{N}}$ of time changes that fulfills (B.1) for $(\alpha_n,\beta_n,\gamma_n)$ and (α,β,γ). We claim that this same sequence fulfills (B.1) for the last d components of $\varphi_k(\alpha_n,\beta_n,\gamma_n)$ and $\varphi_k(\alpha,\beta,\gamma)$. Then Lemma B.1 yields $\varphi_k(\alpha_n,\beta_n,\gamma_n) \xrightarrow{S} \varphi_k(\alpha,\beta,\gamma)$. We write for short $\tau_k^n = \tau_k(\alpha_n)$ and $\tau_k = \tau_k(\alpha)$, and let $N \in \mathbb{N}$, then

$$\begin{aligned}
&\sup_{t\le N}\left|\sum_{\ell\in I}\gamma_n^\ell(\lambda_n(t))\ \mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}} - \sum_{\ell\in I}\gamma^\ell(t)\ \mathbb{1}_{\{\alpha(\tau_k)=\ell\}}\right| \\
&\quad\le \sum_{\ell\in I}\sup_{t\le N}\left|\gamma_n^\ell(\lambda_n(t))\ \mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}} - \gamma^\ell(t)\ \mathbb{1}_{\{\alpha(\tau_k)=\ell\}}\right| \\
&\quad\le \sum_{\ell\in I}\sup_{t\le N}\left(\left|\gamma_n^\ell(\lambda_n(t)) - \gamma^\ell(t)\right| \mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}} + \left|\gamma^\ell(t)\right|\left|\mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}} - \mathbb{1}_{\{\alpha(\tau_k)=\ell\}}\right|\right)
\end{aligned}$$

by the triangular inequality. Now $(\alpha_n,\beta_n,\gamma_n) \xrightarrow{S} (\alpha,\beta,\gamma)$ implies $\alpha_n(\tau_k^n) \to \alpha(\tau_k)$ by Proposition B.5, so since α_n and α take only values in the finite set I, there exists $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$ we have $\alpha(\tau_k) = \ell$ if and only if $\alpha_n(\tau_k^n) = \ell$. (Here we need the fact that we have convergence $\alpha_n \xrightarrow{S} \alpha$ in the Skorokhod topology on $\mathbb{D}(I \times E' \times (\mathbb{R}^d)^m)$ rather than on $\mathbb{D}(\mathbb{R} \times E' \times (\mathbb{R}^d)^m)$.) Thus $\mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}} - \mathbb{1}_{\{\alpha(\tau_k)=\ell\}}$ vanishes for all $n \ge n_0$, and since $|\mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}}| \le 1$, we have

$$\sup_{t\le N}\left|\sum_{\ell\in I}\gamma_n^\ell(\lambda_n(t))\ \mathbb{1}_{\{\alpha_n(\tau_k^n)=\ell\}} - \sum_{\ell\in I}\gamma^\ell(t)\ \mathbb{1}_{\{\alpha(\tau_k)=\ell\}}\right| \le \sum_{\ell\in I}\sup_{t\le N}\left|\gamma_n^\ell(\lambda_n(t)) - \gamma^\ell(t)\right|$$

for n sufficiently large, which tends to 0 since $\gamma_n^\ell \xrightarrow{S} \gamma^\ell$ for all $\ell \in I$ with the above $(\lambda_n)_{n\in\mathbb{N}}$ by assumption. □

With the help of the next proposition it is possible to show convergence of solutions of stochastic differential equations driven by point processes like in Theorem 1.53 and Proposition 5.4.

Proposition B.8 *Let* $d, d' \in \mathbb{N}$ *and let* $f_k\colon (\mathbb{Z}^d)^k \to \mathbb{Z}^{d'}$ *for* $k \in \mathbb{N}$*. Define for some fixed* $y_0 \in \mathbb{Z}^d$ *the mapping* $\varphi\colon \mathbb{D}(\mathbb{Z}^d) \to \mathbb{D}(\mathbb{Z}^{d'})$ *by*

$$\varphi(\alpha)(t) = \sum_{k=0}^{\infty}\left(y_0 + f_k\Big(\Delta\alpha(\tau_1(\alpha)),\ldots,\Delta\alpha(\tau_k(\alpha))\Big)\right)\mathbb{1}_{[\tau_k,\tau_{k+1})}.$$

Then the mapping $\Phi\colon \mathbb{D}(\mathbb{Z}^d) \to \mathbb{D}(\mathbb{Z}^d \times \mathbb{Z}^{d'})$*, given by* $\Phi(\alpha)(t) = (\alpha(t), \varphi(\alpha)(t))$ *is continuous for the Skorokhod topology on* $\mathbb{D}(\mathbb{Z}^d)$*.*

In order to prove Proposition B.8, we need to show that $\alpha_n \xrightarrow{S} \alpha$ implies $\Phi(\alpha_n) \xrightarrow{S} \Phi(\alpha)$ for $\alpha_n, \alpha \in \mathbb{D}(\mathbb{Z}^d)$. The idea is to construct a sequence of time changes $(\hat{\lambda}_n)_{n\in\mathbb{N}}$ such that $\hat{\lambda}_n$ shifts the jump times of α onto those of α_n. From Proposition B.5, we know that for all $k \in \mathbb{N}$ the mapping $\alpha \mapsto \tau_k(\alpha)$ is continuous for the Skorokhod topology, however it fails to be uniformly continuous in k, so we need to stop shifting jumps after a certain number of jumps.

To keep notation short we denote the jump times of α and α_n by $\tau_k := \tau_k(\alpha)$ and $\tau_k^n := \tau_k(\alpha_n)$. We also define for $b > 0$ and $n \in \mathbb{N}$

$$M(n,b) := \sup\left\{ M \in \mathbb{N} \,\middle|\, \sup_{1\le k\le M} |\tau_k - \tau_k^\ell| < b \text{ for all } \ell \ge n \right\}, \tag{B.3}$$

and the sequence $(b_n)_{n\in\mathbb{N}}$ by

$$\begin{aligned} b_1 &= 1 \\ b_{n+1} &= \begin{cases} b_n & \text{if } M\left(n+1, \frac{b_n}{2}\right) \le M(n,b_n), \\ \frac{b_n}{2} & \text{if } M\left(n+1, \frac{b_n}{2}\right) > M(n,b_n). \end{cases} \end{aligned} \tag{B.4}$$

Lemma B.9 *For $\alpha_n \xrightarrow{S} \alpha$ and $M(n,b)$ and b_n as defined above we have*

a) $\lim_{n\to\infty} b_n = 0$,

b) $\sup_{1\le k\le M(n,b_n)} |\tau_k - \tau_k^n| \le b_n$ *for all* $n \in \mathbb{N}$,

c) $\lim_{n\to\infty} M(n,b_n) = \infty$.

Proof. The mapping $\alpha \mapsto \tau_k$ is continuous for the Skorokhod topology, so $\tau_k^n \to \tau_k$ for all k and therefore $\lim_{n\to\infty} M(n,b) = \infty$ for fixed $b > 0$. Thus for fixed $n \in \mathbb{N}$ there exists ℓ such that $M(n+\ell, \frac{b}{2}) > M(n,b)$. Then by construction of b_n, there exists ℓ such that $b_{n+\ell} = \frac{b_n}{2}$. So by induction we have that for all n there exists n_0 such that $b_{n_0} = 2^{-n}$, and since b_n is decreasing, we have $b_\ell \le 2^{-n}$ for all $\ell \ge n_0$, which shows a).

Now b) is satisfied by construction of $M(n,b)$, so it remains to show c). However by construction of b_n for all $n \in \mathbb{N}$ there exists k such that $M(n+k, b_{n+k}) > M(n,b_n)$, so $\lim_{n\to\infty} M(n,b_n) = \infty$ since $M(n,b_n)$ only takes values in $\mathbb{N}$. □

Lemma B.10 *Let $n \in \mathbb{N}$ and $\alpha, \alpha_n \in \mathbb{D}$. Denote by τ_k and τ_k^n the k-th jump time of α and α_n, respectively. For $M(n,b)$ and b_n as above let*

$$\begin{aligned} \hat{\lambda}_n(t) &= \sum_{k=1}^{M(n,b_n)} \left(\tau_{k-1}^n + \frac{\tau_k^n - \tau_{k-1}^n}{\tau_k - \tau_{k-1}} (t - \tau_{k-1}) \right) \mathbb{1}_{[\tau_{k-1},\tau_k)}(t) \\ &\quad + \left(t + \tau_{M(n,b_n)}^n - \tau_{M(n,b_n)} \right) \mathbb{1}_{[\tau_{M(n,b_n)},\infty)}(t). \end{aligned} \tag{B.5}$$

Then $\hat{\lambda}_n(\tau_k) = \tau_k^n$ for all $k \le M(n,b_n)$, i.e., $\hat{\lambda}_n$ shifts the first $M(n,b_n)$ jump times of α onto those of α_n. Furthermore $(\hat{\lambda}_n)_{n\in\mathbb{N}} \subseteq \Lambda$, and $(\hat{\lambda}_n)_{n\in\mathbb{N}}$ satisfies (B.1 i).

Proof. By construction $\hat{\lambda}_n$ is strictly increasing, continuous, and satisfies $\hat{\lambda}_n(0) = 0$ and $\lim_{t\to\infty} \hat{\lambda}_n(t) = \infty$ for all $n \in \mathbb{N}$. It is also clear that $\hat{\lambda}_n(\tau_k) = \tau_k^n$ for all $k \leq M(n, b_n)$, so it remains to show that (B.1 i) is satisfied by $(\hat{\lambda}_n)_{n\in\mathbb{N}}$. Now for fixed k and $\tau_{k-1} \leq t < \tau_k$ we have

$$\begin{aligned}\left|\tau_{k-1}^n + \frac{\tau_k^n - \tau_{k-1}^n}{\tau_k - \tau_{k-1}}(t - \tau_{k-1}) - t\right| &= \left|\left(\tau_{k-1}^n - \tau_{k-1}\right)\frac{\tau_k - t}{\tau_k - \tau_{k-1}} + \left(\tau_k^n - \tau_k\right)\frac{t - \tau_{k-1}}{\tau_k - \tau_{k-1}}\right| \\ &\leq \left|\tau_{k-1}^n - \tau_{k-1}\right|\left|\frac{\tau_k - t}{\tau_k - \tau_{k-1}}\right| + \left|\tau_k^n - \tau_k\right|\left|\frac{t - \tau_{k-1}}{\tau_k - \tau_{k-1}}\right| \\ &\leq \left|\tau_{k-1}^n - \tau_{k-1}\right| + \left|\tau_k^n - \tau_k\right|,\end{aligned}$$

since $\left|\frac{\tau_k - t}{\tau_k - \tau_{k-1}}\right| \leq 1$ and $\left|\frac{t - \tau_{k-1}}{\tau_k - \tau_{k-1}}\right| \leq 1$ for $\tau_{k-1} \leq t < \tau_k$. So for fixed n we have

$$\begin{aligned}\sup_{t\geq 0}|\hat{\lambda}_n(t) - t| &\leq \sup_{t\geq 0}\left(\sum_{k=1}^{M(n,b_n)}\left|\tau_{k-1}^n + \frac{\tau_k^n - \tau_{k-1}^n}{\tau_k - \tau_{k-1}}(t - \tau_{k-1}) - t\right|\mathbb{1}_{[\tau_{k-1},\tau_k)}(t)\right. \\ &\qquad \left. + \left|\tau_{M(n,b_n)}^n - \tau_{M(n,b_n)}\right|\mathbb{1}_{[\tau_{M(n,b_n)},\infty)}(t)\right) \\ &\leq 2\sup_{k\leq M(n,b_n)}\left|\tau_k^n - \tau_k\right|.\end{aligned}$$

Lemma B.9 b) then yields that $\sup_{t\geq 0}|\hat{\lambda}_n(t) - t| \leq 2b_n$, which tends to 0 as $n \to \infty$ by Lemma B.9 a), hence $(\hat{\lambda}_n)_{n\in\mathbb{N}}$ satisfies condition (B.1 i). $\square$

Proof of Proposition B.8. Let $\alpha_n \xrightarrow{S} \alpha$ in $\mathbb{D}(\mathbb{Z}^d)$. We need to show that this implies $(\alpha_n, \varphi(\alpha_n)) \xrightarrow{S} (\alpha, \varphi(\alpha))$ in $\mathbb{D}(\mathbb{Z}^d \times \mathbb{Z}^{d'})$, and by Lemma B.1 it suffices to show $\alpha_n \xrightarrow{S} \alpha$ and $\varphi(\alpha_n) \xrightarrow{S} \varphi(\alpha)$ separately as long as (B.1 ii) holds with the same sequence of time changes. We claim that $(\hat{\lambda}_n)_{n\in\mathbb{N}}$ from (B.5) is suitable for this purpose. By Lemma B.10 $(\hat{\lambda}_n)_{n\in\mathbb{N}}$ satisfies (B.1 i), hence it remains to show that (B.1 ii) holds with $(\hat{\lambda}_n)_{n\in\mathbb{N}}$ for both α_n, α and $\varphi(\alpha_n), \varphi(\alpha)$. So we fix $N \in \mathbb{N}$, and we need to show that $\sup_{t\leq N}|\alpha_n \circ \hat{\lambda}_n(t) - \alpha(t)| \xrightarrow{n\to\infty} 0$ and $\sup_{t\leq N}|\varphi(\alpha_n) \circ \hat{\lambda}_n(t) - \varphi(\alpha)(t)| \xrightarrow{n\to\infty} 0$.

Concerning the first convergence, recall that $\hat{\lambda}_n$ shifts the first $M(n, b_n)$ jump times of α onto those of α_n. Let k_N be the number of jumps of α on $[0, N]$ and choose n_1 big enough so that $M(n, b_n) \geq k_N$ for all $n \geq n_1$. This is always possible since $M(n, b_n) \uparrow \infty$ as $n \to \infty$ by Lemma B.9 c). Now for fixed $t \leq N$ and $n \geq n_1$ there exists $k_t \leq M(n, b_n)$ so that $t \in [\tau_{k_t}, \tau_{k_t+1})$, and by construction of $\hat{\lambda}_n$ we also have $\hat{\lambda}_n(t) \in [\tau_{k_t}^n, \tau_{k_t+1}^n)$. Note that α takes values in $\mathbb{Z}^d$, so it is constant between jumps, and therefore

$$|\alpha_n \circ \hat{\lambda}_n(t) - \alpha(t)| = |\alpha_n(\tau_{k_t}^n) - \alpha(\tau_{k_t})|.$$

Now $\alpha_n \xrightarrow{S} \alpha$ implies $\alpha_n(\tau_k^n) \to \alpha(\tau_k)$ for all $k \in \mathbb{N}$ (cf. Jacod and Shiryaev (1987), Proposition VI.2.7), so there exists $n_2 \in \mathbb{N}$ such that for all $n \geq n_2$ we have $\alpha(\tau_k) = \alpha_n(\tau_k^n)$ for all $k \le k_N$. (Here we need the fact that we have convergence $\alpha_n \xrightarrow{S} \alpha$ in the Skorokhod topology on $\mathbb{D}(\mathbb{R}^m \times I)$ rather than on $\mathbb{D}(\mathbb{R}^{m+1})$.) So since α has only finitely many jumps on $[0, N]$, we have

$$\sup_{t \le N} |\alpha_n \circ \hat{\lambda}_n(t) - \alpha(t)| = 0$$

for n sufficiently large, hence (B.1 ii) is fulfilled for α_n, α with $(\hat{\lambda}_n)_{n\in\mathbb{N}}$.

Regarding $\sup_{t \le N} |\varphi(\alpha_n) \circ \hat{\lambda}_n(t) - \varphi(\alpha)(t)| \overset{n\to\infty}{\longrightarrow} 0$ we choose n_1 and n_2 as above so that $M(n, b_n) \ge k_N$ and $\Delta\alpha_n(\tau_k^n) = \Delta\alpha(\tau_k)$ for all $k \le k_N$. Then for all $n \ge n_1 \vee n_2$ and arbitrary $t \le N$, to which we associate a unique $k_t \le M(n, b_n)$ so that $t \in [\tau_{k_t}, \tau_{k_t+1})$, we have

$$\begin{aligned}
&|\varphi(\alpha_n) \circ \hat{\lambda}_n(t) - \varphi(\alpha)(t)| \\
&= \; |f_k(\Delta\alpha_n(\tau_1^n), \ldots, \Delta\alpha_n(\tau_k^n)) - f_k(\Delta\alpha(\tau_1), \ldots, \Delta\alpha(\tau_k))| \; \mathbb{1}_{[\tau_{k_t}, \tau_{k_t+1})}(t) \\
&= \; 0,
\end{aligned}$$

so that $\sup_{t \le N} |\varphi(\alpha_n) \circ \hat{\lambda}_n(t) - \varphi(\alpha)(t)| = 0$ for $n \ge n_1 \vee n_2$, which shows (B.1 ii) for $\varphi(\alpha_n), \varphi(\alpha)$ with $(\hat{\lambda}_n)_{n\in\mathbb{N}}$. □

Appendix C

Some Technical Results

On Part II

Here we give the proofs of some technical results needed in Part II on truncation functions and properties of the convex functions f and g. Recall that in Lemma 2.13 we have defined the functions $f \geq g\colon [0,\infty) \to [0,\infty)$ by $f(y) = y \log y - (y-1)$ and $g(y) = (1-\sqrt{y})^2$. For the following lemmas let U be a $d \times d$-matrix, let K be a measure which integrates $1 \wedge |x|^2$ and define $\nu(ds, dx) = dsK(dx)$.

Lemma C.1 *If h_1, h_2 are two truncation functions, then*

$$\int_{\mathbb{R}^d} |h_1(x) - h_2(x)| \; K(dx) < \infty.$$

Proof. Since h_1, h_2 are bounded and $h_1(x) - h_2(x) = 0$ for $|x| \leq \varepsilon$ or $|x| \geq \frac{1}{\varepsilon}$ for some $0 < \varepsilon < 1$, we have $|h_1(x) - h_2(x)| \leq \text{const.}(1 \wedge |x|^2)$ and thus

$$\int_{\mathbb{R}^d} |h_1(x) - h_2(x)| \; K(dx) \leq \text{const.} \int_{\mathbb{R}^d} 1 \wedge |x|^2 \; K(dx),$$

which is finite by assumption. □

Lemma C.2 *Let $z, y \geq 0$. Then for all $\alpha > 0$*

$$zy \leq \frac{1}{\alpha}\left(f(y) + e^{\alpha z} - 1\right).$$

Proof. Let $a > 0$. A simple calculation shows that

$$f(ay) = af(y) + y \; a \log a - (a-1),$$

so we get

$$\begin{aligned} f(e^{-\alpha z} y) &= e^{-\alpha z} f(y) + y e^{-\alpha z} \log e^{-\alpha z} - (e^{-\alpha z} - 1) \\ &= e^{-\alpha z}\left(f(y) - \alpha z y + e^{\alpha z} - 1\right). \end{aligned}$$

Since $f \geq 0$, this yields $\alpha z y \leq f(y) + e^{\alpha z} - 1$, hence the result. □

Lemma C.3 *Let Y be a nonnegative predictable function. Then*

$$|Uh(x) - h(Ux)|Y * \nu_t \leq \text{const.}(t + f(Y) * \nu_t)$$

for all $t \geq 0$.

Proof. Since h is a truncation function, we have $h(x) = x$ for $|x| \leq \varepsilon'$ for some $\varepsilon' > 0$, and since U defines a linear (and thus uniformly continuous) function on $\mathbb{R}^d$, there exists some $\varepsilon'' > 0$ such that $|Ux| \leq \varepsilon'$ and thus $h(Ux) = Ux$ for all $|x| \leq \varepsilon''$. This yields $Uh(x) - h(Ux) = 0$ for $|x| < \varepsilon := \varepsilon' \wedge \varepsilon''$. Furthermore h is bounded, so that

$$|Uh(x) - h(Ux)| \leq |Uh(x)| + |h(Ux)| \leq \text{const.}$$

for some finite constant. Altogether we have

$$|Uh(x) - h(Ux)| \leq \text{const.}\, \mathbb{1}_{\{|x| \geq \varepsilon\}}$$

for some $\varepsilon > 0$. Thus

$$\begin{aligned} |Uh(x) - h(Ux)|Y * \nu_t &\leq \text{const.}\, \big(\mathbb{1}_{\{|x|\geq\varepsilon\}} Y(s,x)\big) * \nu_t \\ &\leq \text{const.} \left(f(Y) * \nu_t + t \int_{\mathbb{R}^d} (e^\alpha - 1) \mathbb{1}_{\{|x|\geq\varepsilon\}}\, K(dx) \right) \end{aligned}$$

from Lemma C.2. Note that $(e^\alpha - 1)\mathbb{1}_{\{|x|\geq\varepsilon\}} \leq \text{const.}(1 \wedge |x|^2)$, so that the integral is finite. $\square$

Lemma C.4 *Let $0 \leq b \leq 1$ and $a > 0$, then*

$$b|a-1| \leq f(a) + b^2.$$

Proof. The claim is obvious for $a = 1$, so let $a \neq 1$ and define $p_a(b) := b^2 - |a-1|b + f(a)$. We show that $p_a(b) \geq 0$ for $0 \leq b \leq 1$ for any choice of $a > 0$, $a \neq 1$. Since $p_a(0) = f(a) > 0$, it suffices to show that all zeros of p_a, if they exist, are greater than 1. The zeros of p_a are given by

$$b_{1,2} = \frac{|a-1| \pm \sqrt{|a-1|^2 - 4f(a)}}{2},$$

and we claim that if the radicand is nonnegative, then $b_{1/2} > 1$. (If the radicand is negative, p_a does not have any real zeros at all and is thus positive everywhere.) So let $|a-1|^2 - 4f(a) \geq 0$. Then $b_{1/2} > 1$ is equivalent to $||a-1|-2| > \sqrt{|a-1|^2 - 4f(a)}$ or $|a-1| - 1 < f(a)$, which is clear for $a < 2$ since then the left-hand side is negative, whereas the right-hand side is always nonnegative. So let $a \geq 2$, thus $|a-1| = a-1$, and we show that $k(a) := f(a) - (a-2)$ is strictly positive for $a \geq 2$. Now $k'(a) = \log a - 1$, and $k''(a) = \frac{1}{a}$, so there is a global minimum of k at $a = e$, and $k(e) = 3 - e > 0$, hence the claim. $\square$

Lemma C.5 *Let h be a truncation function and Y be a nonnegative predictable function. Then*

$$|h\ (Y-1)| * \nu_t \leq \text{const.}(t + f(Y) * \nu_t)$$

for all $t \geq 0$.

Proof. Note that $|h(x)| \leq \text{const.}\left(|x|\mathbb{1}_{\{|x|\leq 1\}} + \mathbb{1}_{\{|x|>1\}}\right)$, so that we get

$$\begin{aligned} |h\ (Y-1)| * \nu_t &\leq \text{const.}\left(\left(|x|\mathbb{1}_{\{|x|\leq 1\}}\ |Y-1|\right) * \nu_t + \left(\mathbb{1}_{\{|x|>1\}}|Y-1|\right) * \nu_t\right) \\ &\leq \text{const.}\left(\mathbb{1}_{\{|x|\leq 1\}}\left(f(Y) + |x|^2\right) * \nu_t + \left(\mathbb{1}_{\{|x|>1\}}\left(f(Y(x)) + 1\right)\right) * \nu_t\right) \\ &= \text{const.}\left(f(Y(x)) * \nu_t + (1 \wedge |x|^2) * \nu_t\right), \end{aligned}$$

where in the second inequality we have used Lemma C.4 with $(b,a) = (|x|, Y(s,x))$ and $(b,a) = (1, Y(s,x))$, respectively. Now

$$(1 \wedge |x|^2) * \nu_t = \int_0^t \int_{\mathbb{R}^d} (1 \wedge |x|^2)\ K(dx)ds = t \int_{\mathbb{R}^d} (1 \wedge |x|^2)\ K(dx),$$

and the claim follows from the fact that K integrates $1 \wedge |x|^2$. □

We close with the following specification of Jensen's inequality.

Lemma C.6 *Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X \in \mathcal{L}^1(P)$. Let $f\colon \mathbb{R} \to \mathbb{R}$ be a strictly convex function. Then $f(E[X]) \leq E[f(X)]$ and*

$$f(E[X]) = E[f(X)] \quad \textit{if and only if} \quad X = E[X]\ P\text{-a.s.}$$

Proof. The first statement is classical. For the second statement sufficiency is immediate, so regarding necessity we start as in the classical proof of Jensen's inequality. For all $x_0 \in \mathbb{R}$ there exists $\lambda(x_0)$ such that

$$\begin{aligned} f(x) &> f(x_0) + (x - x_0)\ \lambda(x_0) \quad \text{if and only if } x \neq x_0, \\ f(x) &= f(x_0) + (x - x_0)\ \lambda(x_0) \quad \text{if and only if } x = x_0. \end{aligned}$$

Set $x_0 = E[X]$ and $x = X$, this yields

$$\begin{aligned} &f(X) > f(E[X]) + (X - E[X])\ \lambda(E[X]) \text{ on } \{X \neq E[X]\} \\ &f(X) = f(E[X]) \text{ on } \{X = E[X]\}. \end{aligned}$$

So suppose $P[X \neq E[X]] > 0$, then

$$\begin{aligned} E[f(X)] &= E[f(X)\mathbb{1}_{\{X\neq E[X]\}}] + E[f(X)\mathbb{1}_{\{X=E[X]\}}] \\ &> f(E[X])\ P[X \neq E[X]] + \lambda(E[X])\ E[(X - E[X])\ \mathbb{1}_{\{X\neq E[X]\}}] + \\ &\quad + f(E[X])\ P[X = E[X]] \\ &= f(E[X]). \end{aligned}$$

Note that $E[(X - E[X])\ \mathbb{1}_{\{X\neq E[X]\}}] = 0$. □

On Part III

In Part III, more specifically in Section 6.2, we recursively define certain functions, where we usually start with a C^1 function. By the following lemma we see that under certain conditions the C^1 property is preserved under these recursions.

Lemma C.7 *Let $g\colon \mathbb{R}^d \times \mathbb{R}^{d'} \to \mathbb{R}^+$ be a bounded C^1-function and let Y be a random variable taking values in $\mathbb{R}^{d'}$. Denote by $\nabla_x g$ the vector of partial derivatives of g with respect to the first d coordinates. If there exists some $Z \in \mathcal{L}^1(P)$ such that $|\nabla_x g(x,Y)| \le Z$ for all $x \in \mathbb{R}^d$, then $w\colon \mathbb{R}^d \to \mathbb{R}^+$, defined by*

$$w(x) = E[g(x,Y)],$$

is a bounded C^1-function and its gradient is given by $\nabla w(x) = E[\nabla_x g(x,Y)]$.

Proof. It is immediate that w is bounded if g is bounded. If furthermore g is C^1, then for $\varepsilon \in \mathbb{R}^d$

$$g(x+\varepsilon, y) = g(x,y) + \nabla_x g(x,y)\ \varepsilon + r(x,\varepsilon,y) \tag{C.1}$$

with $\lim_{\varepsilon\to 0} \frac{r(x,\varepsilon,y)}{|\varepsilon|} = 0$. Thus

$$\begin{aligned} w(x+\varepsilon) &= E[g(x+\varepsilon, Y)] \\ &= E[g(x,Y)] + E[\nabla_x g(x,Y)\ \varepsilon] + E[r(x,\varepsilon,Y)] \\ &= w(x) + E[\nabla_x g(x,Y)]\varepsilon + E[r(x,\varepsilon,Y)]. \end{aligned}$$

The claim is established if we show $\lim_{\varepsilon\to 0} E\left[\frac{r(x,\varepsilon,Y)}{|\varepsilon|}\right] = 0$. However by (C.1) and a Taylor expansion of first order (which is nothing else than the mean value theorem in $\mathbb{R}^d$) we have

$$\begin{aligned} \frac{|r(x,\varepsilon,y)|}{|\varepsilon|} &\le |\nabla_x g(x,y)| + \frac{|g(x+\varepsilon,y) - g(x,y)|}{|\varepsilon|} \\ &\le |\nabla_x g(x,y)| + |\nabla_x g(x+\vartheta\varepsilon, y)| \end{aligned}$$

for some $\vartheta \in [0,1]$. Thus by assumption $\frac{r(x,\varepsilon,Y)}{|\varepsilon|} \le 2Z$, so dominated convergence yields the result. $\square$

In Section 6.2 g has some special structure, and we state the following result. For a matrix-valued function $\Psi\colon \mathbb{R}^d \to \mathbb{R}^{d\times r}$ we denote the matrix $\left(\frac{\partial \Psi^{ij}}{\partial x^k}(x)\right)_{i\in\{1,\dots,d\}, j\in\{1,\dots,r\}}$ by $\frac{\partial \Psi}{\partial x^k}(x)$.

Lemma C.8 *Let $Y = (Y^\ell)_{\ell\in\{1,\dots,r\}}$ be a square-integrable $\mathbb{R}^d$-valued random variable with $E[Y] = 0$ and Y^ℓ, Y^j uncorrelated for $\ell \ne j$. Let $f^*\colon \mathbb{R}^d \to \mathbb{R}$ and $\Psi\colon \mathbb{R}^d \to \mathbb{R}^{d\times r}$ be measurable and define $f(x) = E[f^*(x + \Psi(x)\ Y)]$. If f^* and Ψ are C^1 with $|\nabla f^*| \le c$ and $|\frac{\partial \Psi^{j\ell}}{\partial x^i}| \le \bar\Psi$ for some positive constants c and $\bar\Psi$, then*

$$\max_{1\le k\le d} \left| \frac{\partial f(x)}{\partial x^k} \right| \le c\sqrt{1 + d\bar\Psi^2 E\left[|Y|^2\right]}$$

for all $x \in \mathbb{R}^d$.

Proof. Let $g(x,y) = f^*(x + \Psi(x)y)$. Then by Lemma C.7 and the chain rule we have for $k \in \{1,\dots,d\}$

$$\begin{aligned}
\frac{\partial f}{\partial x^k}(x) &= E[\nabla_x g(x,Y)] \\
&= E\left[\left(\nabla f^*(x+\Psi(x)\ Y)\right)^{\mathrm{tr}} \frac{\partial \Psi}{\partial x^k}(x)\ Y + \frac{\partial f^*}{\partial x^k}(x+\Psi(x)\ Y)\right] \\
&= E[v^{\mathrm{tr}}(w_k + z_k)],
\end{aligned}$$

where v, w_k, z_k are $\mathbb{R}^d$-valued random variables with

$$v = \nabla f^*(x+\Psi(x)\ Y), \quad w_k^i = \delta_{ki}, \quad z_k^i = \sum_{j=1}^r \frac{\partial \Psi^{ij}}{\partial x^k}(x)\ Y^j, \quad i \in \{1,\dots,d\}.$$

So by the Cauchy-Schwarz inequality we get

$$\left(\frac{\partial f}{\partial x^k}(x)\right)^2 \le E\left[|v|^2\right] E\left[|w_k + z_k|^2\right],$$

where $E\left[|v|^2\right] \le c^2$ by assumption, and

$$|w_k + z_k|^2 = \sum_{i=1}^d (w_k^i + z_k^i)^2 = 1 + 2z_k^k + |z_k|^2.$$

So since $E[z_k^k] = E\left[\sum_{j=1}^r \frac{\partial \Psi^{kj}}{\partial x^k}(x)\ Y^j\right] = 0$, and

$$\begin{aligned}
E[|z_k|^2] &= E\left[\sum_{i=1}^d \left(\sum_{j=1}^r \frac{\partial \Psi^{ij}}{\partial x^k}(x)\ Y^j\right)^2\right] \\
&= \sum_{i=1}^d E\left[\sum_{\ell,j=1}^r \frac{\partial \Psi^{ij}}{\partial x^k}(x) \frac{\partial \Psi^{i\ell}}{\partial x^k}(x)\ Y^j\ Y^\ell\right] \\
&\le d\bar{\Psi}^2 E\left[|Y|^2\right]
\end{aligned}$$

because Y^j, Y^ℓ are uncorrelated, we get

$$\left(\frac{\partial f}{\partial x^k}(x)\right)^2 \le c^2\left(1 + d\bar{\Psi}^2 E\left[|Y|^2\right]\right).$$

This does not depend on k, so that the claim follows. $\square$

We next show the

Proof of Lemma 5.13 c). We need to show that $p'^{n\ell\pi} \le \frac{c}{n^2}$ where $p'^{n\ell\pi} = 1 - \sum_{j=1}^m p'^{n\ell j}$ for $p'^{n\ell j}$ from (5.24). Notice that for $m \le 2$ we have $\sum_{j=1}^m p'^{n\ell j} = 1$ by construction, so that in particular in this case η^n never leaves $\{1,\ldots,m\}$. Thus $p'^{n\ell\pi}=0$, and there is nothing to show.

Now let $m\geq 3$. We show that $\sum_{j=1}^m p'^{n\ell j}\geq 1-\frac{c''}{n^2}$; then the result follows from the fact that $p'^{n\ell\pi}=1-\sum_{j=1}^m p'^{n\ell j}$. Now $\sum_{j=1}^m p'^{n\ell j}=P'^n\left[\bigcup_{j=1}^m A_k^{n\ell j}\right]$, and $\bigcup_{j=1}^m A_k^{n\ell j}$ is the event that $\eta^n_{t_k}$ does not leave $\{1,\ldots,m\}$ in one step given $\eta^n_{t_{k-1}}=\ell$. This happens at least if $\zeta_k^{n\ell i}=0$ for all $i\in\{1,\ldots,m\}$, or if only one $\zeta_k^{n\ell i}=1$, whereas the others are 0, or if $\zeta_k^{n\ell\ell}=1$, $\zeta_k^{n\ell i}=1$ for only one $i\neq\ell$ and $\zeta_k^{n\ell r}=0$ for all $r\notin\{\ell,j\}$. Thus

$$\begin{aligned}
P'^n\left[\bigcup_{j=1}^m A_k^{n\ell j}\right] &\geq P'^n\left[\bigcap_{i=1}^m\left\{\zeta_k^{n\ell i}=0\right\}\right]+P'^n\left[\bigcup_{i=1}^m\left(\{\zeta_k^{n\ell i}=1\}\cap\bigcap_{\substack{r=1\\ r\neq i}}^m\{\zeta_k^{n\ell r}=0\}\right)\right]\\
&\quad+P'^n\left[\{\zeta_k^{n\ell\ell}=1\}\cap\left(\bigcup_{\substack{i=1\\ i\neq\ell}}^m\left(\{\zeta_k^{n\ell i}=1\}\cap\bigcap_{\substack{r=1\\ r\notin\{\ell,j\}}}^m\{\zeta_k^{n\ell r}=0\}\right)\right)\right].
\end{aligned}$$

Therefore we have by the independence of $\zeta_k^{n\ell i}$ under P'^n

$$\begin{aligned}
P'^n\left[\bigcup_{j=1}^m A_k^{n\ell j}\right] &\geq \left(1-\frac{T}{n}\right)^m+m\frac{T}{n}\left(1-\frac{T}{n}\right)^{m-1}+(m-1)\left(\frac{T}{n}\right)^2\left(1-\frac{T}{n}\right)^{m-2}\\
&= \left(1-\frac{T}{n}\right)^{m-2}\left(\left(1-\frac{T}{n}\right)^2+m\frac{T}{n}\left(1-\frac{T}{n}\right)+(m-1)\left(\frac{T}{n}\right)^2\right)\\
&= \left(1-\frac{T}{n}\right)^{m-2}\left(1+(m-2)\frac{T}{n}\right)\\
&= \left(1-(m-2)\frac{T}{n}+\begin{pmatrix}m-2\\2\end{pmatrix}\left(\frac{T}{n}\right)^2+\mathcal{O}\left(\frac{1}{n^3}\right)\right)\left(1+(m-2)\frac{T}{n}\right)\\
&= 1-(m-2)^2\left(\frac{T}{n}\right)^2+\begin{pmatrix}m-2\\2\end{pmatrix}\left(\frac{T}{n}\right)^2+\mathcal{O}\left(\frac{1}{n^3}\right)\\
&= 1-\frac{1}{2}(m-1)(m-2)\left(\frac{T}{n}\right)^2+\mathcal{O}\left(\frac{1}{n^3}\right),
\end{aligned}$$

and therefore $\sum_{j=1}^m p'^{n\ell j}=P'^n\left[\bigcup_{j=1}^m A_k^{n\ell j}\right]\geq 1-\frac{c''}{n^2}$, for some appropriate constant c''. □

We close with two fairly simple but useful results concerning conditional expectations.

Lemma C.9 *Let X and Y be $\mathbb{R}^d$-valued random variables and $|X|\leq c$ P-a.s. Then for any sub-σ-algebra $\mathcal{G}\subseteq\mathcal{F}$*

$$E\left[(XY-E[XY|\mathcal{G}])^2\Big|\mathcal{G}\right]\leq c^2E\left[|Y|^2\big|\mathcal{G}\right]\quad P\text{-a.s.}$$

Proof. This is immediate since

$$E\left[(XY - E[XY|\mathcal{G}])^2 \middle| \mathcal{G}\right] = E\left[(XY)^2 \middle| \mathcal{G}\right] - \left(E\left[XY \middle| \mathcal{G}\right]\right)^2 \leq c^2 E\left[|Y|^2 \middle| \mathcal{G}\right],$$

because $E\left[(XY)^2 \middle| \mathcal{G}\right] \leq E\left[|X|^2|Y|^2 \middle| \mathcal{G}\right]$ and $\left(E\left[XY \middle| \mathcal{G}\right]\right)^2 \geq 0$. $\square$

Lemma C.10 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra, $\mathcal{G}^0$ its P-completion and let X be an integrable $\mathbb{R}$-valued random variable. Then*

$$E[X|\mathcal{G}] = E[X|\mathcal{G}^0].$$

Proof. Since $\mathcal{G} \subseteq \mathcal{G}^0$ it is clear that $E[X|\mathcal{G}]$ is $\mathcal{G}^0$-measurable and it remains to show that $E[X\mathbb{1}_G] = E[E[X|\mathcal{G}]\mathbb{1}_G]$ for all $G \in \mathcal{G}^0$. So let $G \in \mathcal{G}^0$. Then there exists $G' \in \mathcal{G}$ with $\mathbb{1}_G = \mathbb{1}_{G'}$ P-a.s., so that the assertion follows from the $\mathcal{G}$-measurability of $E[X|\mathcal{G}]$. $\square$

Appendix D

Comments on Fujiwara and Miyahara (2003)

Fujiwara and Miyahara (2003) introduce a one-dimensional model with finite time horizon, where the price process of some tradable asset is modeled by an *exponential Lévy process*, i.e. $X = e^{\tilde{L}}$, where $\tilde{L}$ is a P-Lévy process. (In fact they call this a geometric Lévy process, however we think that this terminology is misleading.) By Example 2.5 this model can be transferred to our model, where the stock prices are given by stochastic exponentials of Lévy processes, i.e. what we call geometric Lévy process. Fujiwara and Miyahara (2003) construct a martingale measure Q^* and show that it has minimal entropy with an admittedly nice stopping argument. The model of geometric Lévy processes has the advantage that one does not have to assume that the jumps of the Lévy process are bounded from below by -1 in order to get positivity of X, however the drawback is that, in contrast to our approach, in general X is no longer a local martingale if $\tilde{L}$ is and vice versa. We compare their results with ours and list certain points which are not presented in full perspicuousness. For the sake of better readability we do so in our notation and under the assumption of zero interest rate (i.e. $r = 0$). At the end of Appendix D we provide a list of notational differences.

Let us quickly review their model. Let $\tilde{L}$ be a P-Lévy process with finite time horizon $[0, T]$ and P-Lévy characteristics $(\tilde{b}, \tilde{c}, \tilde{K})$ relative to the canonical truncation function $h_0(x) = \mathbb{1}_{\{|x| \leq 1\}}$ and define $X = e^{\tilde{L}}$.

In their Section 2 Fujiwara and Miyahara (2003) state that X may be written as the stochastic exponential of a (different) Lévy process, i.e. $e^{\tilde{L}_t} = \mathcal{E}(L)_t$, for a P-Lévy process L with P-Lévy characteristics (b, c, K), where

$$\begin{aligned} b &= \tilde{b} + \frac{1}{2}\tilde{c} + \int (h_0(J(y)) - h_0(y))\ \tilde{K}(dy), \\ c &= \tilde{c}, \\ K(dx) &= \tilde{K} \circ J^{-1}(dx), \end{aligned}$$

for $J(y) = e^y - 1$, cf. to our Example 2.5.

For their main result, Theorem 3.1, Fujiwara and Miyahara (2003) state the following condition.

(C) There exists a constant $u^* \in \mathbb{R}$ which satisfies

$$
\begin{aligned}
&(i)_{\mathrm{FM}} \qquad \int_{\{y>1\}} e^y e^{u^*(e^y-1)}\ \tilde{K}(dy) < \infty \\
&(ii)_{\mathrm{FM}} \qquad \tilde{b} + \left(\frac{1}{2} + u^*\right)\tilde{c} + \int_{\{|y|\leq 1\}} \left((e^y-1)e^{u^*(e^y-1)} - y\right)\ \tilde{K}(dy) + \\
&\qquad\qquad + \int_{\{|y|>1\}} (e^y-1)e^{u^*(e^y-1)}\ \tilde{K}(dy) = 0.
\end{aligned}
$$

In their Theorem 3.1 Fujiwara and Miyahara (2003) assume condition (C) and

(1) construct the candidate measure Q^* by

$$
\frac{dQ^*}{dP} = \frac{e^{u^* L_T}}{E_P\left[e^{u^* L_T}\right]},
$$

(2) claim that $\tilde{L}$ is a Lévy process under Q^*,

(3) claim that Q^* has minimal entropy among all absolutely continuous local martingale measures for X and give an explicit calculation of $I(Q^*|P)$.

Note that condition (C) in Fujiwara and Miyahara (2003) is actually stronger than our condition on the existence of the Esscher martingale measure $Q^* = Q^{u^*}$ (relative to L) from our Theorem 4.7 in a one-dimensional setting. If we translate our condition to the setting of a exponential Lévy process via our Example 2.5, we get

$$
\begin{aligned}
&(i)_{\mathrm{E}} \qquad \int_{\mathbb{R}} \left|(e^y-1)e^{u^*(e^y-1)} - h(J(y))\right|\ \tilde{K}(dy) < \infty \\
&(ii)_{\mathrm{E}} \qquad \tilde{b} + \left(\frac{1}{2} + u^*\right)\tilde{c} + \int_{\mathbb{R}} \left((e^y-1)e^{u^*(e^y-1)} - h(y)\right)\ \tilde{K}(dy) = 0.
\end{aligned}
$$

Fujiwara and Miyahara (2003) use the canonical truncation function $h_0(x) = \mathbb{1}_{\{|x|\leq 1\}}$, so $(ii)_{\mathrm{FM}}$ and $(ii)_{\mathrm{E}}$ for $h = h_0$ are clearly equivalent. Furthermore by the triangle inequality and the $\tilde{K}$-integrability of $h(J(y)) - h(y)$, condition $(i)_{\mathrm{E}}$ is equivalent to

$$
(i)'_{\mathrm{E}} \qquad \int_{\mathbb{R}} \left|(e^y-1)e^{u^*(e^y-1)} - h(y)\right|\ \tilde{K}(dy) < \infty.
$$

If we set $h = h_0$ then we see that condition $(i)'$ further reduces to

$$
(i)''_{\mathrm{E}} \qquad \int_{y>1} (e^y-1)e^{u^*(e^y-1)}\ \tilde{K}(dy) < \infty.
$$

In fact by a Taylor expansion around 0 we see that the constant and the linear term of $(e^y-1)e^{u^*(e^y-1)}-y$ vanish, so that

$$(e^y-1)e^{u^*(e^y-1)}-y=\mathcal{O}(y^2)$$

as $y\stackrel{n\to\infty}{\longrightarrow}0$, so that

$$\int_{|y|<1}\left|(e^y-1)e^{u^*(e^y-1)}-h(y)\right|\ \tilde{K}(dy)<\infty,$$

since $\tilde{K}$ integrates $1\wedge|y|^2$. On $\{y<-1\}$ we have $-1<e^y-1<0$, so that for $u^*\geq 0$ we have $0<e^{u^*(e^y-1)}<1$, whereas for $u^*<0$ we get $1<e^{u^*(e^y-1)}<e^{-u^*}$, so altogether

$$\left|(e^y-1)e^{u^*(e^y-1)}\right|<1\vee e^{-u^*}$$

on $\{y<-1\}$. So $(i)_{\mathrm{FM}}$ clearly implies $(i)''_{\mathrm{E}}$, since for $y>1$ we have

$$(e^y-1)e^{u^*(e^y-1)}<e^ye^{u^*(e^y-1)}.$$

In the following we list the points where the reasoning of Fujiwara and Miyahara (2003) seems unclear to us. We also suggest how our results can give answers to these questions.

1. Note that Q^*, defined by

$$\frac{dQ^*}{dP}=\frac{e^{u^*L_T}}{E_P\left[e^{u^*L_T}\right]}$$

is an Esscher martingale measure for L but *not* an Esscher measure for $\tilde{L}$. In fact, in Shiryaev (1999), Theorem VII.3.4 and also in Bühlmann, Delbaen, Embrechts and Shiryaev (1996) it is shown that an Esscher measure $Q^{\tilde{u}^*}$ for $\tilde{L}$ is a martingale measure for $X=e^{\tilde{L}}$, if

$$(a)\qquad \int_{\mathbb{R}}\left|(e^y-1)e^{\tilde{u}^*y}-h(e^y-1)\right|\ \tilde{K}(dy)<\infty$$

$$(b)\qquad \tilde{b}+\left(\frac{1}{2}+\tilde{u}^*\right)\tilde{c}+\int_{\mathbb{R}}\left((e^y-1)e^{\tilde{u}^*y}-h(y)\right)\ \tilde{K}(dy)=0.$$

Note the difference in the exponent compared to our conditions $(i)_{\mathrm{E}}$ and $(ii)_{\mathrm{E}}$. Indeed, the Girsanov quantities of $Q^{\tilde{u}^*}$ with respect to P *relative to* L are in this case

$$\begin{aligned}\beta &= \tilde{u}^*\\ Y(y) &= (1+y)^{\tilde{u}^*}\end{aligned}$$

(cf. Example 2.5). A consequence of this is that the proof of X being a local Q^*-martingale is unclear in Fujiwara and Miyahara (2003) because their martingale condition (3.11) in Remark 3.2 (2) is not the one as cited from Shiryaev (1999) or Bühlmann, Delbaen, Embrechts and Shiryaev (1996).

However with the martingale condition from Theorem 4.7 rewritten in terms of Lévy characteristics of $\tilde{L}$ as seen in $(ii)_{\mathrm{E}}$, their condition (3.11) holds. The reasoning then should be: X is a strictly positive local Q^*-martingale, hence a supermartingale, and since it has constant expectation it is a true Q^*-martingale. ◇

2. As in our Section 4.3, Fujiwara and Miyahara (2003) claim that the measure Q^* with

$$Z_t^* = \left.\frac{dQ^*}{dP}\right|_{\mathcal{F}_t} = \frac{e^{u^* L_t}}{E_P\left[e^{u^* L_t}\right]} \tag{D.1}$$

is the minimal entropy martingale measure. They claim that $Z^* = \mathcal{E}(N^*)$ for

$$N_t^* = u^* \sqrt{c} W_t + \left(e^{u^*(e^y - 1)} - 1\right) * (\mu^{\tilde{L}} - \tilde{\nu}^P)_t,$$

where W is a standard Brownian motion under P and $\tilde{\nu}^P$ denotes the P-compensator of $\mu^{\tilde{L}}$. However it is not clear that the stochastic integral with respect to $\mu^{\tilde{L}} - \tilde{\nu}^P$ on the right hand side is well-defined. By their definition of stochastic integral with respect to a compensated random measure, taken from Ikeda and Watanabe (1989), Section II.3, a predictable function W is integrable with respect to $\mu^{\tilde{L}} - \tilde{\nu}^P$ if either $|W| * \tilde{\nu}_t^P$ is P-integrable, or if $W^2 * \tilde{\nu}^P$ is locally P-integrable (then $W * \mu^{\tilde{L}} - \tilde{\nu}^P$ is a locally square-integrable martingale by our Theorem 1.22 a)). Neither of these assertions seem to be directly implied by condition $(i)_{\mathrm{FM}}$ as they claim.

However there is the following way out: By construction, Q^* is an Esscher measure for L, so the Girsanov quantities are given by $(u^*, e^{u^* x})$ (cf. our Proposition 4.5), and by our Corollary 2.9

$$\left.\frac{dQ^*}{dP}\right|_{\mathcal{F}_t} = \mathcal{E}\left(u^* \sqrt{c} W + \left(e^{u^* x} - 1\right) * (\mu^L - \nu^P)\right)_t,$$

where W is P-Brownian motion. In particular $e^{u^* x} - 1$ is integrable with respect to $(\mu^L - \nu^P)$, so (recall $\nu(ds, dx) = ds K(dx)$)

$$\int_{\mathbb{R}} \left(1 - \sqrt{e^{u^* x}}\right)^2 K(dx) < \infty \tag{D.2}$$

by our Theorem 1.22 c). Now $K(dx) = \tilde{K} \circ J^{-1}(dy)$, so (D.2) implies

$$\int_{\mathbb{R}} \left(1 - \sqrt{e^{u^*(e^y - 1)}}\right)^2 \tilde{K}(dy) < \infty,$$

thus $e^{u^*(e^y - 1)} - 1$ is in fact integrable with respect to $\mu^{\tilde{L}} - \tilde{\nu}^P$. ◇

3. The proof of the optimality of Q^* in Fujiwara and Miyahara (2003) relies of the following stopping argument: Let $Q \in \mathcal{Q}_a(X)$ be an arbitrary martingale measure for X. Then as in our Proposition 4.6 L is a local Q-martingale, hence there exists an increasing sequence of stopping times $(\tau_n)_{n \in \mathbb{N}}$ with $\tau_n \uparrow T$ Q-a.s. such that L^{τ_n} is a Q-martingale. Furthermore $\mathcal{F}_{\tau_n} \subseteq \mathcal{F}_T$, so $I_{\mathcal{F}_{\tau_n}}(Q|P) \leq I_T(Q|P)$. By construction of Q^* as an Esscher measure, we have for $t \in [0, T]$

$$Z_t^* = u^* L_t - \Psi(u^*) t$$

by our Theorem A.7, where

$$\Psi(w) = b^{\mathrm{tr}}w + \frac{1}{2}w^{\mathrm{tr}}cw + \int_{\mathbb{R}^d} \left(e^{w^{\mathrm{tr}}x} - 1 - w^{\mathrm{tr}}x1\!\!1_{\{|x|\le 1\}}\right) K(dx)$$

is well-defined whenever $\int_{\{|x|>1\}} e^{w^{\mathrm{tr}}x} K(dx) < \infty$. Then along the localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ we have

$$\log Z^*_{\tau_n} = u^* L_{\tau_n} - \Psi(u^*)\tau_n$$

Now L^{τ_n} is a Q-martingale, and $\tau_n \le T$, so $\log Z^*_{\tau_n}$ is Q-integrable, and thus (cf. the Remark after Theorem 2.1 in Csiszár (1975))

$$I_{\mathcal{F}_{\tau_n}}(Q|P) = I_{\mathcal{F}_{\tau_n}}(Q|Q^*) + E_Q\left[\log \left.\frac{dQ^*}{dP}\right|_{\mathcal{F}_{\tau_n}}\right] \ge E_Q\left[\log \left.\frac{dQ^*}{dP}\right|_{\mathcal{F}_{\tau_n}}\right] = E_Q\left[\log Z^*_{\tau_n}\right].$$

For the first equality we need the finiteness of either $I_{\mathcal{F}_{\tau_n}}(Q|P)$ or $I_{\mathcal{F}_{\tau_n}}(Q|Q^*)$, therefore we will have to choose $Q \in \mathcal{Q}_f$. Actually we want to show that Q^* is optimal in $\mathcal{Q}_a$, but any $Q \in \mathcal{Q}_a \cap \mathcal{Q}_f^c$ has infinite relative entropy, so if $\mathcal{Q}_f \ne \emptyset$ it suffices to show that Q^* is optimal in $\mathcal{Q}_f$.

Now altogether we have

$$\begin{aligned} I_T(Q|P) &\ge E_Q\left[\log Z^*_{\tau_n}\right] \\ &= E_Q\left[u^* L_T^{\tau_n} - \Psi(u^*)\tau_n\right] \\ &= -\Psi(u^*)E_Q[\tau_n], \end{aligned}$$

since L^{τ_n} is a Q-martingale. Now with $n \to \infty$ we get by monotone convergence

(D.3) $$I_T(Q|P) \ge -\Psi(u^*)T.$$

It remains to show that $I_T(Q^*|P) = -\Psi(u^*)T$. This is already stated in the proof of our Lemma 4.3; Fujiwara and Miyahara (2003) claim that L may be written as

$$L_t = cW^*_t + (e^y - 1) * (\mu^{\tilde{L}} - \tilde{\nu}^{Q^*})_t$$

where W^* is Q^*-Brownian motion and $\tilde{\nu}^{Q^*}$ denotes the Q^*-compensator of $\mu^{\tilde{L}}$. To us it is not clear how to deduce this form of L from the Lévy-Itô decomposition of $\tilde{L}$ under Q^* neither why the stochastic integral with respect to $\mu^{\tilde{L}} - \tilde{\nu}^{Q^*}$ is well-defined. ◇

4. Once this is established Fujiwara and Miyahara (2003) state that

(D.4) $$I(Q^*|P) = E_{Q^*}\left[\log Z^*_T\right] = E_{Q^*}\left[u^* L_T - \Psi(u^*)T\right] = -\Psi(u^*)T,$$

which together with (D.3) yields the optimality of Q^*.

Apparently Fujiwara and Miyahara (2003) claim in (D.4) that $E_{Q^*}[L_T] = 0$ which is of course true because L is a local martingale under Q^* *and* a Lévy process since Q^* is an Esscher measure for L (cf. our Proposition 4.2), and thus a true Q^*-martingale. However this fact is stated nowhere in their article. ◇

List of Notational Differences

Fujiwara and Miyahara	Esche	explanation
$S = e^X = \mathcal{E}(\hat{X})$		stock price process
$X, \hat{X}$		Lévy processes
$\tilde{S}$	X	discounted stock price process
$\tilde{R} = \hat{X} - rt$	L	return process, driving Lévy process
		(*stochastic* logarithm of $\tilde{S}$, respectively X)
N_p		jump measure of X
N_q	μ^L	jump measure of $\tilde{R}$ respectively L
$\hat{N}_p(ds, dx)$		P-compensator of N_p
$= ds\ \nu(dx)$		
$\hat{N}_q(ds, dx)$	$\nu^P(ds, dx)$	P-compensator of N_q respectively μ^L
$= ds\ \mu(dx)$	$= ds\ K(dx)$	
$\tilde{N}_p = N_p - \hat{N}_p$		compensated measure
$\tilde{N}_q = N_q - \hat{N}_q$	$\mu^L - \nu^P$	compensated measure
P^*	Q^*	candidate for the optimal measure
β_*	u^*	parameter for Esscher martingale measure

Bibliography

Antonelli, F., and A. Kohatsu-Higa (2000): "Filtration stability of backward SDEs," *Stochastic Analysis and Applications*, 18(1), 11–37.

Becherer, D. (2001): "Rational hedging and valuation with utility-based preferences," Ph.D. thesis, Technical University Berlin,
available on `http://edocs.tu-berlin.de/diss/2001/becherer_dirk.htm`.

Becherer, D., and M. Schweizer (2003): "Classical solutions to reaction-diffusion systems for hedging with interacting Itô and point processes," Preprint, Imperial College London.

Bellini, F., and M. Frittelli (2002): "On the existence of minimax martingale measures," *Mathematical Finance*, 12(1), 1–21.

Bertoin, J. (1996): *Lévy Processes.* Cambridge University Press, Cambridge.

Billingsley, P. (1968): *Convergence of Probability Measures.* John Wiley & Sons, New York.

Black, F., and M. Scholes (1973): "The pricing of options and corporate liabilities," *Journal of Political Economy*, 81(3), 637–654.

Bouchard, B., and N. Touzi (2002): "Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations," Preprint, CREST ENSAE Paris.

Brémaud, P. (1981): *Point Processes and Queues. Martingale Dynamics.* Springer-Verlag, Berlin.

Briand, P., B. Delyon and J. Mémin (2001): "Donsker-type theorem for BSDEs," *Electronic Communications in Probability*, 6, 1–14.

Bühlmann, H., F. Delbaen, P. Embrechts and A. N. Shiryaev (1996): "No-arbitrage, change of measure and conditional Esscher transforms," *CWI Quarterly*, 9(4), 291–317.

Chan, T. (1999): "Pricing contingent claims on stocks driven by Lévy processes," *Annals of Applied Probability*, 9(2), 504–528.

Coquet, F., J. Mémin and L. Słominski (2001): "On weak convergence of filtrations," in *Séminaire de Probabilités XXXV, Lecture Notes in Mathematics 1755*, 306–328. Springer-Verlag, Berlin.

Cox, J. C., S. A. Ross and M. Rubinstein (1979): "Option pricing: a simplified approach," *Journal of Financial Economics*, 7(3), 229–263.

Csiszár, I. (1975): "I-divergence geometry of probability distributions and minimization problems," *Annals of Probability*, 3, 146–158.

DELBAEN, F., P. GRANDITS, T. RHEINLÄNDER, D. SAMPERI, M. SCHWEIZER, AND C. STRICKER (2002): "Exponential hedging and entropic penalties," *Mathematical Finance*, 12(2), 99–123.

DELLACHERIE, C., AND P.-A. MEYER (1980): *Probabilités et Potentiel. Chapitres V à VIII: Théorie des Martingales*. Hermann, Paris.

DOUGLAS, J. JR., J. MA AND P. PROTTER (1996): "Numerical methods for forward-backward stochastic differential equations," *Annals of Applied Probability*, 6(3), 940–968.

DUFFIE, D., AND L. G. EPSTEIN (1992): "Stochastic differential utility," *Econometrica*, 60(2), 353–394.

DUFFIE, D., AND P. PROTTER (1992): "From discrete- to continuous-time finance: weak convergence of the financial gain process," *Mathematical Finance*, 2(1), 1–15.

DURRETT, R. (1991): *Probability. Theory and Examples*. Wadsworth & Brooks/Cole, Pacific Grove.

ESCHE, F., AND M. SCHWEIZER (2003): "Minimal entropy preserves the Lévy property: How and why," Preprint, Technical University Berlin.

ETHIER, S. N., AND T. G. KURTZ (1986): *Markov Processes. Characterization and Convergence*. John Wiley & Sons, New York.

FOLDES, L. (1990): "Certainty equivalence in the continuous-time portfolio-cum-saving model," in *Applied Stochastic Analysis, Stochastics Monographs 5*, 343–387. Gordon and Breach, London.

——— (1991): "Optimal sure portfolio plans," *Mathematical Finance*, 1(2), 15–55.

FÖLLMER, H., AND M. SCHWEIZER (1990): "Hedging of contingent claims under incomplete information," in *Applied Stochastic Analysis, Stochastics Monographs 5*, 389–414. Gordon and Breach, London.

FRITTELLI, M. (2000): "The minimal entropy martingale measure and the valuation problem in incomplete markets," *Mathematical Finance*, 10(1), 39–52.

FUJIWARA, T., AND Y. MIYAHARA (2003): "The minimal entropy martingale measures for geometric Lévy processes," *Finance and Stochastics*, 7(4), 509–531.

GERBER, H. U., AND E. S. SHIU (1994): "Option pricing by Esscher transforms," *Transactions of the Society of Actuaries*, XLVI, 99–191.

——— (1996): "Actuarial bridges to dynamic hedging and option pricing," *Insurance: Mathematics & Economics*, 18(3), 183–218.

GOLL, T., AND J. KALLSEN (2000): "Optimal portfolios for logarithmic utility," *Stochastic Processes and their Applications*, 89(1), 31–48.

GOLL, T., AND L. RÜSCHENDORF (2001): "Minimax and minimal distance martingale measures and their relationship to portfolio optimization," *Finance and Stochastics*, 5(4), 557–581.

GRANDITS, P., AND T. RHEINLÄNDER (2002): "On the minimal entropy martingale measure," *Annals of Probability*, 31(3), 1003–1038.

HE, S.-W., J.-G. WANG AND J.-A. YAN (1992): *Semimartingale Theory and Stochastic Calculus*. Science Press, Beijing.

HOOVER, D. N. (1991): "Convergence in distribution and Skorokhod convergence for the general theory of processes," *Probability Theory and Related Fields*, 89(3), 239–259.

Hubalek, F., and T. Hudetz (1998): "Convergence of option prices under minimum entropy martingale measure," unpublished manuscript.

Hubalek, F., and W. Schachermayer (1998): "When does convergence of asset price processes imply convergence of option prices?," *Mathematical Finance*, 8(4), 385–403.

Ikeda, N., and S. Watanabe (1989): *Stochastic Differential Equations and Diffusion Processes.* North-Holland, Amsterdam.

Jacod, J., S. Méléard and P. Protter (2000): "Explicit form and robustness of martingale representations," *Annals of Probability*, 28(4), 1747–1780.

Jacod, J., J. Mémin and M. Métivier (1983): "On tightness and stopping times," *Stochastic Processes and their Applications*, 14, 109–146.

Jacod, J., and A. N. Shiryaev (1987): *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin.

Jakubowski, A., J. Mémin and G. Pagès (1989): "Convergence en loi des suites d'intégrales stochastiques sur l'espace D^1 de Skorokhod," *Probability Theory and Related Fields*, 81(1), 111–137.

Kabanov, Yu., and C. Stricker (2001): "On equivalent martingale measures with bounded densities," in *Séminaire de Probabilités XXXV, Lecture Notes in Mathematics 1755*, 139–148. Springer-Verlag, Berlin.

——— (2002): "On the optimal portfolio for the exponential utility maximization: remarks to the six-author paper," *Mathematical Finance*, 12(2), 125–134.

Kallenberg, O. (1997): *Foundations of Modern Probability.* Springer-Verlag, Berlin.

Kallsen, J. (2000): "Optimal portfolios for exponential Lévy processes," *Mathematical Methods of Operations Research*, 51(3), 357–374.

Kallsen, J., and A. N. Shiryaev (2002): "The cumulant process and Esscher's change of measure," *Finance and Stochastics*, 6(4), 397–428.

Kullback, S. (1959): *Information Theory and Statistics.* John Wiley & Sons, New York.

Kurtz, T. G., and P. E. Protter (1991): "Weak limit theorems for stochastic integrals and stochastic differential equations," *Annals of Probability*, 19(3), 1035–1070.

——— (1996): "Weak convergence of stochastic integrals and differential equations," in *Lecture Notes in Mathematics 1627*, 1–41. Springer-Verlag, Berlin.

Lepingle, D., and J. Mémin (1978): "Sur l'intégrabilité uniforme des martingales exponentielles," *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 42, 175–203.

Lesne, J.-P., J.-L. Prigent and O. Scaillet (2000): "Convergence of discrete time option pricing models under stochastic interest rates," *Finance and Stochastics*, 4(1), 81–93.

Ma, J., P. Protter, J. San Martín and S. Torres (2002): "Numerical method for backward stochastic differential equations," *Annals of Applied Probability*, 12(1), 302–316.

Ma, J., P. Protter and J. Yong (1994): "Solving forward-backward stochastic differential equations explicitly – a four step scheme," *Probability Theory and Related Fields*, 98(3), 339–359.

Miyahara, Y. (2001): "[Geometric Lévy processes & MEMM] Pricing model and related estimation problems," *Asia-Pacific Financial Markets*, 8(1), 15–60.

PRIGENT, J.-L. (1999): "Incomplete markets: Convergence of option values under the minimal martingale measure," *Advances in Applied Probability*, 31(4), 1058–1077.

PROTTER, P. (1990): *Stochastic Integration and Differential Equations. A New Approach.* Springer-Verlag, Berlin.

RACHEV, S., AND L. RÜSCHENDORF (1994): "Models for option prices," *Theory of Probability and its Applications*, 39(1), 120–152.

SATO, K.-I. (1999): *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge.

SCHACHERMAYER, W. (2001): "Optimal investment in incomplete markets when wealth may become negative," *Annals of Applied Probability*, 11(3), 694–734.

SCHWEIZER, M. (1995): "On the minimal martingale measure and the Föllmer-Schweizer decomposition," *Stochastic Analysis and Applications*, 13(5), 573–599.

SHIRYAEV, A. N. (1999): *Essentials of Stochastic Finance.* World Scientific, Singapore.

SŁOMINSKI, L. (1989): "Stability of strong solutions of stochastic differential equations," *Stochastic Processes and their Applications*, 31(2), 173–202.

WONG, E., AND M. ZAKAI (1965): "On the convergence of ordinary integrals to stochastic integrals," *Annals of Mathematics and Statistics*, 36, 1560–1564.

XIA, J., AND J.-A. YAN (2000): "The utility maximization approach to a martingale measure constructed via Esscher transforms," Preprint, Chinese Academy of Sciences, Beijing.